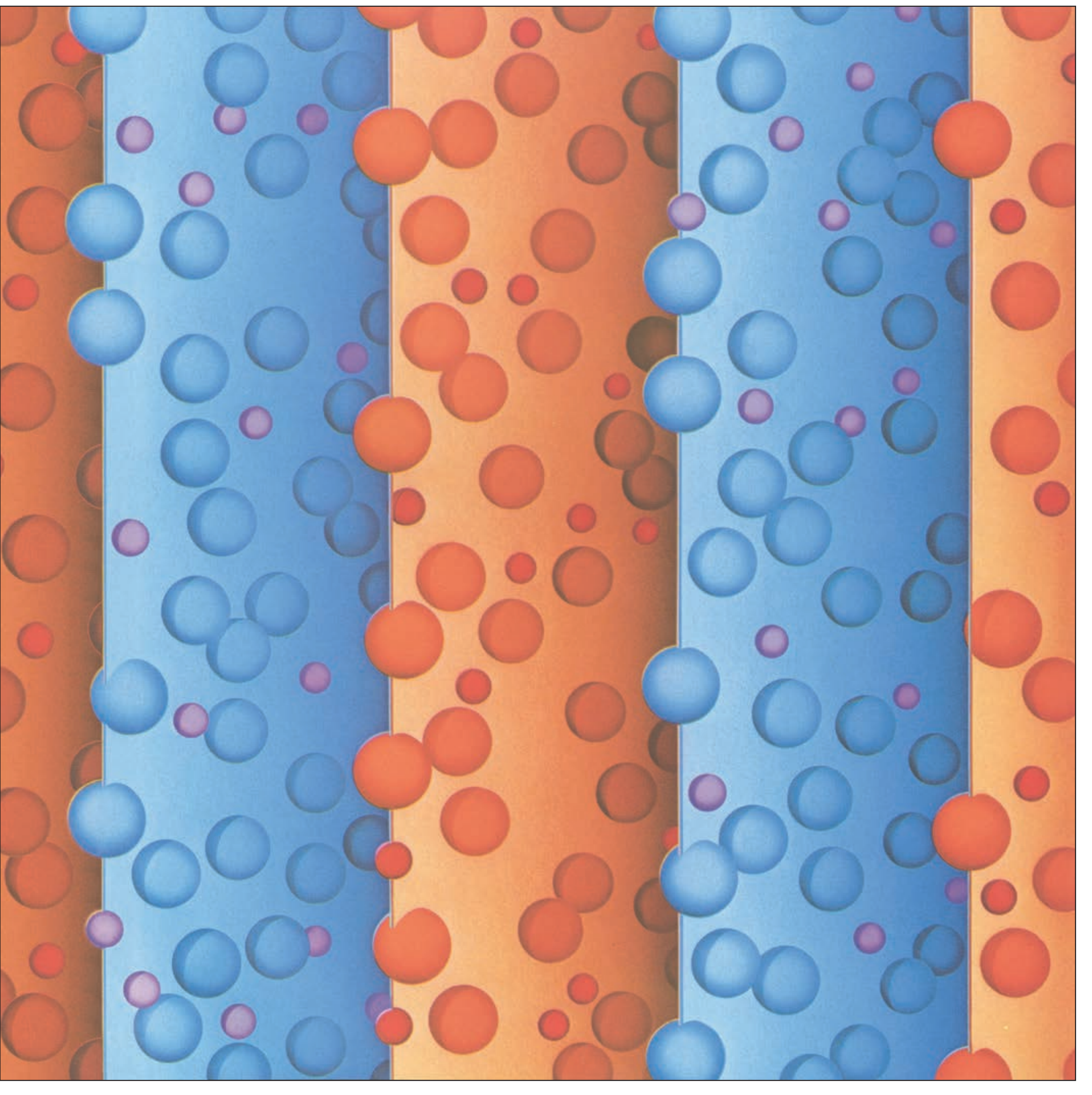
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



SUPERESTRUCTURAS DE SEMICONDUCTORES

Enero 1984
350 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 88
Enero 1984

INVESTIGACION Y
CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

EDITA Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Fernando José Rodilla: *Vulnerabilidad ante un ataque nuclear por sorpresa*; Teodoro Roca: *Centauro-A: la galaxia activa más cercana*; Santiago Torres: *Base molecular del cáncer*; Amando García Rodríguez: *Superestructuras de estado sólido*; Julio Mangas: *Arqueología social de los monumentos megalíticos*; Joandomènec Ros: *Adaptación de un coleóptero a una semilla venenosa*; Andrés Pocoví: *La erupción de Krakatoa*; Luis Bou: *Juegos de ordenador*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad

Leonardo Villena y José Sánchez Real

Libros:

Ramón Pascual, Francisco Ponz y Juan José Acero



LA PORTADA

La figura de la portada representa a una superestructura de estado sólido. Este cristal está formado por capas ultradelgadas de materiales semiconductores y posee unas propiedades electrónicas y ópticas sorprendentes [véase "Superestructuras de estado sólido", por Gottfried H. Döhler]. Está constituido por un semiconductor dispuesto en forma de capas "contaminado" por impurezas. Los planos de color naranja son capas *n*, contaminadas con átomos dadores (*esferas naranja*); estos proporcionan electrones y se convierten en iones positivos. Los planos azules son capas *p*, contaminadas con átomos aceptores (*esferas azules*); los aceptores capturan electrones y se convierten en iones negativos. La aplicación de un voltaje proporciona electrones libres a las capas *n* (*esferas pequeñas de color rojo*) y "agujeros" a las capas *p* (*esferas de color púrpura*).

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:
España:
Un año (12 números): 3.850 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
350 pesetas
Ejemplar atrasado extraordinario:
450 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49 - Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - Madrid-30
Tel. 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.° - Barcelona-34
Teléfono 204 45 83

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Pintura de la portada de George V. Kelvin

Página	Fuente	Página	Fuente
8-14	George V. Kelvin	50-56	Bunji Tagawa
16	Ilil Arbel	57	Ravi Dhar
17	U.S. Air Force	58	Bunji Tagawa
19-28	Ricardo Génova y Salvador Ordóñez	61	Pierre M. Petroff y Arthur C. Gossard, Bell Laboratories
31	Jack O. Burns, Eric D. Feigelson y Ethan J. Schreier	62-68	Allen Beechel
		71	Colin Renfrew, Universidad de Cambridge
32	Jack O. Burns y R. Marcus Price, Universidad de Nuevo México	72	Andrew Tomko
		73	West Air Photography
34	Ian Worpole	74-76	Andrew Tomko
35	Jack O. Burns, Eric D. Feigelson y Ethan J. Schreier	77	Patricia J. Wynne a partir de Alec Daykin
		78	Colin Renfrew
36-37	Ian Worpole	81	James E. Mattler
38-39	Jack O. Burns, Eric D. Feigelson y Ethan J. Schreier	82-87	Patricia J. Wynne
		89	*Album of Krakatau*
		90-98	Todd Pink
40	Ian Worpole	100	Maurice Krafft
42-43	Leonardo Villena	102-104	Edward Bell
44-45	José Sánchez Real	105	Ilil Arbel (*arriba*) Edward Bell (*abajo*)
49	Erika A. Hartwieg y Jonathan A. King, Instituto de Tecnología de Massachusetts	106-108	Edward Bell
		111	Jearl Walker
		112-115	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Napoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

MATTHEW BUNN y KOSTA TSIPIS ("Vulnerabilidad ante un ataque nuclear por sorpresa") se interesan por los efectos de los sistemas de armas y medidas de control de los mismos. Bunn cursa la especialidad de ciencias políticas en el Instituto de Tecnología de Massachusetts (MIT), e investiga encuadrado en el Programa de Ciencia y Tecnología para la Seguridad Internacional de dicho centro superior. Tsipis es co-director del Programa, que en 1977 puso en marcha con Bernard T. Feld. Griego de nacimiento, emigró en 1954 a los Estados Unidos para estudiar ingeniería eléctrica y ciencias físicas, cuyo doctorado recibió por la Universidad de Columbia. Desde 1966 pertenece al departamento de física del MIT.

SALVADOR ORDOÑEZ ("Las salmueras") da clases de petrología (rocas sedimentarias) en la Facultad de Ciencias Geológicas de la Universidad Complutense de Madrid, donde se doctoró, con premio extraordinario, en 1974. Inició su labor investigadora en el tema de recursos naturales en el campo de las "bauxitas" como becario de la Fundación March. Ha llegado al campo de la investigación en salmueras a través del estudio de los yacimientos españoles de sulfato sódico y de la sedimentación salina en lagunas. Sus estudios actuales en este dominio están dirigidos hacia la aplicación de técnicas de isótopos estables a la sedimentación de materiales carbonáticos y salinos.

JACK O. BURNS y R. MARCUS PRICE ("Centauro-A: la galaxia activa más cercana") son, respectivamente, profesor ayudante de astronomía de la Universidad de Nuevo México y profesor de física y astronomía y director de ese departamento en la misma institución. Burns se licenció por la Universidad de Massachusetts en 1974; se trasladó luego a la Universidad de Indiana, donde preparó el doctorado en astronomía. En 1980, después de dos años de trabajo como postgraduado en el Observatorio Nacional de Radioastronomía de los Estados Unidos, marchó a Nuevo México. Además del tema desarrollado en su artículo, se interesa por los supercúmulos galácticos. Price se licenció por la Universidad estatal de Colorado, tras lo cual marchó al continente australiano doctorándose por la Nacional de Australia en 1966. De vuelta a los Estados Unidos se incorporó al Instituto de Tecnología de Massachusetts, que dejó ocho años después para convertirse en el primer director de radioespectros de la National Science Foundation. Fue uno de los descubridores del efecto Faraday, según el cual, el plano de polarización de las ondas de radio procedentes de fuentes extragalácticas gira al atravesar campos magnéticos del espacio interestelar.

ROBERT A. WEINBERG ("Base molecular del cáncer") es profesor de biología del Instituto de Tecnología de Massachusetts, donde pasó sus años de formación. En 1979 se recibió de doctor en biología. Sus investigaciones postdoctorales las realizó en el Instituto Weizmann de Ciencias, en Israel, y en el Instituto Salk de Estudios Biológicos. Volvió al MIT en el año 1972. Consiguió la plaza de profesor en 1982, estando adscrito al Centro para la Investigación del Cáncer y al departamento de biología. Weinberg es también miembro del Instituto Whitehead de Investigaciones Biomédicas.

GOTTFRIED H. DÖHLER ("Superestructuras de estado sólido") trabaja en el Instituto Max Planck de física del estado sólido con sede en Stuttgart. Así resume su curriculum: "Finalizados mis estudios de licenciatura me convertí en un físico teórico, entendiendo por tal quien pertenece a un grupo de teoría. En realidad, cada vez que iniciaba una investigación teórica descubría la existencia de numerosas consecuencias experimentales o de posibles aplicaciones. En consecuencia, todo mi trabajo se veía motivado por aspectos experimentales o de aplicación, lo cual evitó que me perdiera en las numerosas implicaciones de la teoría original. Tal sucedió con mi trabajo sobre los efectos de los campos eléctricos intensos en los semiconductores, así como en mis estudios posteriores sobre la teoría del transporte en los semiconductores amorfos. Al abordar las superestructuras de semiconductores descubrí un campo donde la teoría, la experimentación y las aplicaciones prácticas convergían".

COLIN RENFREW ("Arqueología social de los monumentos megalíticos") es catedrático de arqueología de la Universidad inglesa de Cambridge, donde se formó. Se licenció, en 1962, por el St. John's College, doctorándose en arqueología en 1965. Después de cumplir el servicio militar como oficial de la Royal Air Force, se incorporó a la Universidad de Sheffield para enseñar prehistoria y arqueología. Más tarde ingresaría en el claustro docente de la Universidad de Southampton, para ocupar la cátedra de arqueología y dirigir el departamento.

GERALD A. ROSENTHAL ("Adaptación de un coleóptero a una semilla venenosa") enseña biología y toxicología en la Universidad de Kentucky. Cursó su carrera en la Facultad de Ciencias Ambientales y Forestales de la Universidad estatal de Nueva York en Syracuse. Completó sus estudios en la de Duke, hasta doctorarse en fisiología y bioquímica vegetales. Durante varios años fue investigador postdoctoral en bioquímica del Instituto Nacional de la Salud de los Estados Unidos, incorporándose luego al departamento de biología de la Universidad Case Western Reserve. En 1972, dejó Case por Kentucky. Pasó un año sabático (el de 1979) en la cátedra Lady Davis de entomología agrícola de la Universidad Hebrea de Jerusalén.

PETER FRANCIS y STEPHEN SELF ("La erupción de Krakatoa") son geólogos que estudiaron en el Imperial College of Science and Technology de Londres. Actualmente trabajan en Texas y comparten un mismo interés por los efectos de las erupciones volcánicas en el clima. Francis nació en Zambia y cursó en Inglaterra sus estudios de geología. Asistió al Imperial College como postgraduado, doctorándose en geología en 1969. Desde 1971 es miembro del departamento de Ciencias de la Tierra de la Universidad a Distancia. Self comenzó su carrera en la Universidad inglesa de Leeds. En 1974 se recibió de doctor en geología por el Imperial College, tras lo cual se trasladó a la Universidad Victoria de Wellington, Nueva Zelanda, para recoger datos de erupciones en el núcleo volcánico de Taupo. Fue a los Estados Unidos para trabajar sobre los efectos de las erupciones volcánicas en el clima, subvencionado por la NASA. En 1979 se incorporó a la Universidad estatal de Arizona, abandonándola tres años más tarde para ocupar una plaza en el departamento de geología de la Universidad de Texas en Arlington.

Hace...

José M.ª López Piñero

... *ciento cincuenta años*

Nació en Madrid Felipe Picatoste Rodríguez. Su padre era un ferviente seguidor de las ideas liberales más avanzadas; durante el trienio constitucional había acompañado como miliciano nacional a las Cortes hasta Cádiz y participado en la defensa de la ciudad contra el ejército absolutista francés, sufriendo más tarde las correspondientes penalidades en la "ominosa década". Picatoste fue seguidor desde joven de las ideas de su padre. Cuando, en 1854, los progresistas restablecieron la milicia nacional, fue uno de los primeros en alistarse y en pelear en las calles de Madrid.

En 1854 había terminado ya sus estudios universitarios de leyes y también de ciencias, en la sección científica especial de la Facultad de Filosofía de Madrid. Durante cinco años (1852-57) fue profesor sustituto de matemáticas en el Instituto de San Isidro, mientras que en la parte final del reinado de Isabel II se dedicó principalmente al periodismo, como redactor de *Las Novedades*.

Su gran ocasión se la ofreció la Revolución de 1868. Ruiz Zorrilla lo nombró jefe del negociado central del Ministerio de Fomento, puesto desde el que realizó una importante labor, colaborando activamente, por ejemplo, en los decretos relativos a la libertad de enseñanza y en el planteamiento del problema de las bibliotecas populares. Más tarde fue también director de la *Gaceta de Madrid*, hasta el final del período revolucionario.

La Restauración motivó, como es lógico, un nuevo cambio en su vida. Seguidor al principio de Ruiz Zorrilla, dirigió durante algún tiempo el periódico político *El Manifiesto*. Más tarde, sin embargo, se afilió al partido de Sagasta y volvió a ser jefe del negociado central del Ministerio de Fomento con un gabinete liberal (1883-85). Trabajó también como periodista, principalmente en el *Heraldo de Madrid* y en varias revistas. Por último, en julio de 1890, ingresó en el Cuerpo de Archiveros y Bibliotecarios. Dos años más tarde, el 29 de septiembre, falleció en el mismo Madrid.

La abundantísima producción escrita de Picatoste es tan dispersa como su misma biografía. Aparte de innumerables trabajos periodísticos de carácter político, publicó libros, folletos y artículos de tema literario e histórico, tradujo una novela de Julio Verne, desarrolló una notable labor de divulgación científica en varias revistas y escribió manuales didácticos sobre matemáticas, física y química, historia natural, fotografía, geografía, historia y religión, así como un diccionario francés-español y otro popular de la lengua castellana. En medio de tan variopinto panorama hay cuatro contribuciones que le aseguran un digno lugar en la cultura y la ciencia españolas de su época. La primera es la *Memoria sobre las Bibliotecas Populares* (1870) que presentó a José de Echegaray cuando éste ocupaba la cartera de Fomento. La segunda, sus estudios sobre terminología matemática titulados *Vocabulario matemático y etimológico* (1862) y *El tecnicismo matemático en el Diccionario de la Academia Española* (1873). La tercera, los *Elementos de matemáticas* (1.ª ed. 1860), el mejor de sus compendios didácticos, que fue repetidas veces editado hasta bien avanzado el período de entreguerras. La cuarta es su obra *Apuntes para una Biblioteca Científica Española del siglo XVI* (1891).

Picatoste publicó otros trabajos relacionados con la historia de la ciencia, pero todos de relieve inferior, ya que en su mayoría son artículos de divulgación. Mención aparte merece su comentario al discurso de Echegaray *Historia de las matemáticas puras en nuestra España* (1866), reproducido en la antología que los hermanos García Camarero han dedicado a la llamada "polémica de la ciencia española". Este comentario sitúa a Picatoste entre los seguidores de la postura apologética en la década anterior al momento en el que desencadenó la segunda fase de dicha polémica encabezada, como es sabido, por Manuel de la Revilla y Marcelino Menéndez Pelayo.

El contexto de los *Apuntes para una Biblioteca Científica Española del siglo XVI* no fue, sin embargo, el mero debate ideológico, sino lo que en otro lugar he llamado "esfuerzo bibliográfico de la *etapa intermedia* del siglo XIX" en torno a la tradición científica española. Su obra es precisamente el último título importante de la serie de repertorios en los que se recogieron los materiales históricos españoles correspondientes a diversas áreas científicas. También es el único centrado en una época con un enfoque interdisciplinar "de ciencias exactas, físicas y naturales", de acuerdo con la denominación de la Real Academia correspondiente. El punto de vista de Picatoste se refleja en la introducción de su libro: "Nuestra historia científica no está tan estudiada como la literaria, por muy diversas causas, explicables sin duda, pero que deben desaparecer ante la conveniencia de llevar la verdad al campo de la historia y ante la especie de certamen a que vienen acudiendo desde hace algunos años todas las naciones de Europa, alegando sus méritos en la historia de la ciencia y olvidando el soberbio y positivista axioma de que la ciencia vive del presente y del porvenir y no necesita para nada conocer los errores y los procedimientos del pasado, que han muerto para siempre.

"Por esta razón hallarán los lectores en nuestro libro muchas comparaciones entre el estado de la ciencia española y la extranjera, así como la indicación de algunos puntos, no completamente esclarecidos, pero que servirán indudablemente de base para estudios más profundos, ya que puede decirse somos los primeros en dar a luz noticias sobre ciertos hechos importantes de la historia científica, como sucede, por ejemplo, respecto de la aplicación del vapor, por Juan Escrivano; de la construcción de telescopios, por Rogete, y de otros no menos curiosos o interesantes.

"En el extranjero, salvo honrosas excepciones, se nos trata con injusticia en materia de ciencia, dominando todavía, aunque parezca imposible, cierta prevención que brotó contra nosotros desde que salimos a conquistar o a molestar a Europa, según dicen, y cierto desdén rencoroso con que nos trataron en nuestra decadencia.

"Europa no ha tenido por muchos años más que dos libros de texto en la historia de la ciencia: el de Montucla, que hace curas y aun Papas a nuestros Profesores, y el de Saverien: el primero desconoce por completo nuestra historia, y del segundo baste decir que no nombra a los españoles en la historia de la Navegación.

"Ante tales errores y aun malevolencias, es un deber patriótico volver por la honra de España, poniendo las cosas en su lugar y ayudando con gratitud a los extranjeros, que han comenzado esta tarea tal vez antes que nosotros".

Premiada en el concurso de la Biblioteca Nacional de 1868, la obra no fue publicada hasta veintitrés años más tarde, al final de la vida de Picatoste,

cuando éste era ya segundo jefe del Cuerpo de Archiveros y Bibliotecarios.

En su contenido hay que distinguir tres aspectos. El menos satisfactorio corresponde a la interpretación, ya que se reduce con frecuencia a reproducir acríticamente los tópicos de la postura apologista de la ciencia española. Los resúmenes biográficos son, en general, muy pobres y proceden de los repertorios anteriores o de los datos resultantes de una consulta superficial de las obras de cada autor. Por el contrario, ofrece una información bibliográfica precisa y de primera mano, debido a lo cual sigue siendo una obra de consulta indispensable para el estudioso de la España de la época.

Aunque se apoya en los repertorios anteriores y en algunos trabajos de erudición, la inmensa mayoría de esta información está directamente basada en las fuentes, sobre todo impresas, y también en un notable número de manuscritos. La precisión es la principal virtud que la separa de casi todas las demás obras sobre el tema publicadas en la España del siglo XIX –desde las de Hernández Morejón y Chinchilla hasta el *Inventario* de Menéndez Pelayo o el discurso de Fernández Vallín– que incluyen frecuentes errores o inexactitudes y no distinguen claramente entre una noticia indirecta y la descripción objetiva de una fuente.

Otra cosa es el problema de las lagunas. Resulta sorprendente que falte tan escaso número de obras impresas españolas de la época y las áreas científicas incluidas en el repertorio de primera importancia histórica. Me atrevería a decir que se limita a cuatro ausencias: la *Suma de la Art de arismetica* (1482) de Francesc Sanct Climent, entre los textos matemáticos; la *Summa* (1536) de Pere Olesa y el *Apparatus* (1605) de



1. Felipe Picatoste. Grabado que figura en el volumen de sus **Ultimos escritos** *que se editó poco después de su muerte (1892)*

APUNTES

PARA UNA

BIBLIOTECA CIENTÍFICA ESPAÑOLA

DEL SIGLO XVI

ESTUDIOS BIOGRÁFICOS Y BIBLIOGRÁFICOS DE CIENCIAS EXACTAS

FÍSICAS Y NATURALES Y SUS INMEDIATAS APLICACIONES EN DICHO SIGLO

POR

DON FELIPE PICATOSTE Y RODRÍGUEZ

OBRA PREMIADA POR LA BIBLIOTECA NACIONAL

EN EL CONCURSO PÚBLICO DE 1868

É IMPRESA Á EXPENSAS DEL ESTADO



MADRID

IMPRENTA Y FUNDICIÓN DE MANUEL TELLO

IMPRESOR DE CÁMARA DE S. M.

Don Evaristo, 8

1891

2. Portada de los **Apuntes** *de Felipe Picatoste, obra en la que culminó el esfuerzo de investigación bibliográfica que durante el pasado siglo se realizó en torno a la tradición científica española*

Juan Bautista Villalpando, entre los de tema físico; y el *Arte Separatoria* (1598) de Diego de Santiago, entre los que tratan sobre química. Por supuesto, las omisiones de obras de importancia secundaria son más numerosas, aparte de que no siempre figuran todas las ediciones de las reseñadas, en especial las traducciones a idiomas extranjeros.

Los textos científicos manuscritos de la España del siglo XVI todavía no han sido recogidos de forma sistemática. Por ello, las noticias reunidas por Picatoste continúan teniendo tanto interés, aunque naturalmente sean mucho más incompletas y con más errores o confusiones que las relativas a las obras impresas. Subrayemos únicamente la atribución tradicional del *Islario* de Alonso de Santa Cruz a Andrés García de Céspedes o la curiosa invención de un fantástico "López Rupecisa" sobre la base de un manuscrito del *Liber lucis* del bajomedieval Joan de Paratallada (Johannes Rupescissa). Ello no puede hacer olvidar, por supuesto, las noticias correctas de un elevado número de manuscritos científicos españoles de la época, ni la descripción objetiva del contenido de varios de especial importancia, como los de Cedillo Díaz, Pedro Chacón, Andrés Espinosa, Francisco Hernández, Juan Bautista Labaña, Diego Pérez de Mesa y el que hasta los recientes estudios de García Diego se atribuía a Juanelo Turriano.

Vulnerabilidad ante un ataque nuclear por sorpresa

Se asegura que Estados Unidos necesita nuevos misiles instalados en tierra, dada la vulnerabilidad de los actuales. El análisis de las incertidumbres ante un ataque nuclear parece indicar que se ha exagerado dicha debilidad

Matthew Bunn y Kosta Tsipis

La disuasión de la guerra nuclear exige que las fuerzas nucleares no sean vulnerables a un ataque preventivo. Durante decenios los responsables de la planificación estratégica, tanto en Estados Unidos como en la Unión Soviética, se han preocupado por la posibilidad de un primer golpe nuclear que desarmara y dejara a la víctima del mismo sin capacidad de devolverlo, mediante una específica acción de represalia. Tanto uno como otro país han gastado ingentes recursos en el esfuerzo de mantener la supervivencia de sus fuerzas nucleares estratégicas.

La preocupación por la vulnerabilidad de las fuerzas nucleares instaladas en tierra aumentó con el desarrollo de los vehículos de reentrada múltiple contra objetivos independientes (los MIRV, siglas de *M*ultiple *i*ndependently *t*argetable *v*ehicles) a finales de la década de 1960. Esta tecnología, que Estados Unidos ensayó primero, permite que un misil porte varias ojivas, cada una de ellas capaz de golpear sobre un objetivo diferente. Es decir, cada misil MIRV puede actuar sobre varios misiles del adversario, lo que, al ser susceptible cada uno de éstos de portar varias ojivas, da una doble ventaja al atacante.

Por ello, los avances de la tecnología MIRV soviética han suscitado el temor de que, una vez que Rusia hubiese desarrollado misiles MIRV que combinen precisión, potencia y seguridad en el funcionamiento, posea dicho bloque militar capacidad para destruir en un primer golpe la mayor parte de los misiles norteamericanos instalados en tierra, recurriendo para ello al uso de tan sólo una fracción de las armas de que dispone. Aunque los misiles instalados en tierra constituyen la cuarta parte de las fuerzas nucleares estratégicas norteamericanas –el resto son misiles balísticos instalados en submarinos (SLBM) y bombarderos estratégicos– crece la preocupación de que la vulnerabilidad del componente ICBM de la "tríada" estratégica pueda suponer, para Occidente, un grave problema de seguridad.

En el otoño de 1977, los peores temores de los planificadores estratégicos norteamericanos se hicieron realidad, al iniciar la Unión Soviética una serie de pruebas de un nuevo sistema de orientación con una precisión muy perfeccionada. Los cálculos simplificados indican que, una vez que haya desplegado un número apropiado de estos nuevos MIRV, Rusia dispondrá de capacidad real para destruir en un primer ataque el grueso de la fuerza estadounidense instalada en tierra. Así surgió lo que se ha dado en denominar "la ventana de la vulnerabilidad", concepto que ha dominado durante algunos años el pensamiento estratégico occidental, sirviendo de justificación principal para el desarrollo de una nueva generación de armas estratégicas estadounidenses, entre ellas los misiles MX y Trident II.

Pero los cálculos simples no hacen justicia a las importantes incertidumbres inherentes a toda evaluación de los resultados de un ataque contra silos. La información de los servicios de inteligencia raras veces es absoluta y, cuando el estado mayor de las fuerzas occidentales aborda la incertidumbre del valor real de parámetros tales como la precisión de los misiles balísticos intercontinentales de la otra parte, tiene que elaborar hipótesis que parecen conservadoras desde el punto de vista de quien se defiende. Desgraciadamente, el proceso oculta a menudo la realidad de que cualquier ataque entrañaría también un alto grado de incertidumbre desde el punto de vista del atacante.

Dado el inmenso poder de destrucción de los modernos arsenales nucleares, cualquier ataque nuclear de primer golpe sería como un juego de azar a una escala desconocida en la historia de la humanidad: lo que estaría sobre la balanza sería el futuro de civilizaciones enteras. Es tal la magnitud de las apuestas sobre la mesa, que cualquier incertidumbre sobre los resultados de un ataque semejante servirá de poderosa disuasión. La pregunta que se hace cualquiera al reflexionar sobre cómo planificar el ataque ya no es la de "qué se espera", sino "qué es lo peor que puede ocurrir". Por ello, al evaluar la posibilidad de un ataque, es imprescindible hacer un cálculo comparativo de las incertidumbres que entrañaría, para realizar luego una evaluación conservadora antes del mismo. Pocas veces se efectúa ese cálculo en los Estados Unidos al hacer la evaluación del equilibrio estratégico. Lo que generalmente se presenta al Congreso y a la opinión pública son los resultados de un ataque idealizado, casi sin fallos, ignorando a menudo las incertidumbres inherentes a dicho golpe. Trataremos, en este artículo, de corregir ese extremo.

Los silos que albergan a los modernos ICBM son estructuras subterráneas de hormigón, "reforzadas" para resistir los efectos de explosiones nucleares. Los efectos nucleares que podrían dañar a los ICBM son de muy diverso tipo, pero portavoces de las Fuerzas Aéreas de los Estados Unidos han señalado que, por regla general, la mayor vulnerabilidad de los silos se da frente a la onda de choque de la explosión nuclear. Por ello, la solidez de un silo se expresa generalmente en térmi-

nos de qué sobrepresión generada por la onda de choque se necesitaría para destruirlo, medida en kilogramos por centímetro cuadrado.

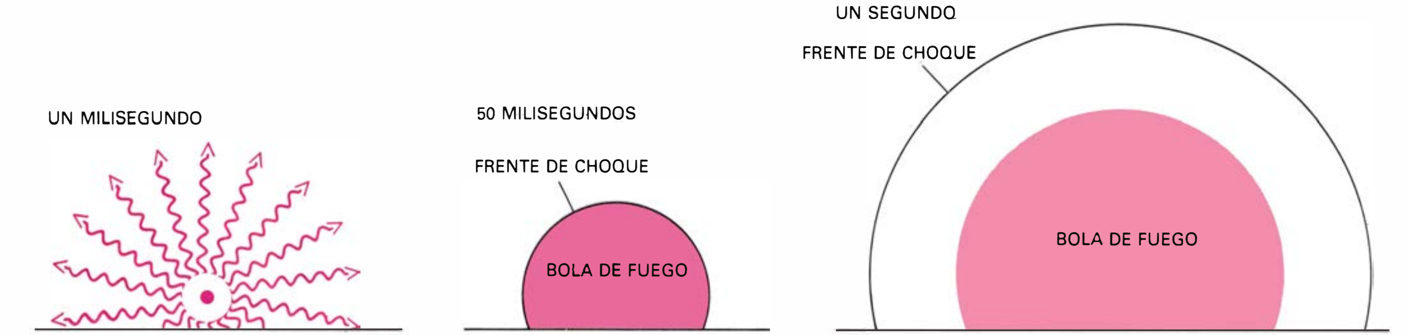
El cálculo más generalizado es que los silos que albergan actualmente los 1000 ICBM Minuteman son capaces de resistir sobrepresiones de hasta 140 kilogramos por centímetro cuadrado. La explosión de un arma de medio megatón, como las que transportan los MIRV soviéticos más precisos, provocaría semejantes sobrepresiones en un rango de 300 metros. Así pues, para destruir un misil Minuteman, habrá que detonar un arma de medio megatón en un radio de 300 metros alrededor del silo. Por ello, aunque la precisión con que caiga el arma no sería especialmente importante en ataques a ciudades, sí tendría decisiva importancia en cualquier ataque contra objetivos reforzados, como son los silos de ICBM.

Los ICBM que transportan esas armas constan de tres partes principales: un cohete (constituido a su vez de varias partes independientes), la carga (que comprende varios RV [*reentry vehicles*], armados con ojivas termonucleares) y el sistema de guía (que dirige el impulso del cohete, a fin de colocar cada RV en la trayectoria adecuada para alcanzar el objetivo asignado). Cuando se lanza uno de esos misiles, el cohete principal sólo desarrolla su función balística entre los tres y los cinco primeros minutos del vuelo. En los minutos siguientes, un cohete más pequeño lleva a cabo los ajustes necesarios de la trayectoria final para poner a cada RV en el camino del objetivo que se pretende. Se sueltan luego los RV, que entran en caída libre en el campo de gravedad de la tierra, sin motor ni guía. Tras media hora de vuelo, los RV retornan a la atmósfera y estallan sobre sus objetivos. En consecuencia, el vuelo del misil puede dividirse en tres fases: la fase propulsada, el vuelo libre y el retorno a la atmósfera.

Hacer caer los RV en el radio de unos cientos de metros de sus objetivos después de recorrer una distancia de 10.000 kilómetros exige un sistema de guía de suma precisión. Las armas estratégicas descansan en una técnica conocida por "guía de inercia", en la que se acude al empleo de giroscopios y acelerómetros para medir las fuerzas que actúan sobre el misil. Pero estos instrumentos inertes no pueden medir la fuerza de la gravedad, dada la equivalencia entre gravedad y aceleración, según describiera Albert Einstein: en un ascensor en caída libre, un acelerómetro señalaría el cero, aunque el ascensor experimentara una aceleración hacia el suelo por influencia de la gravedad. Por ello, para calcular los efectos de la fuerza de atracción de la gravedad de la tierra sobre el desplazamiento del misil, es preciso programar una función de posición en el sistema de guía del misil antes de su lanzamiento. Combinando el modelo de gravedad con las mediciones de determinadas fuerzas por los acelerómetros, el ordenador de guía aplica las leyes de la inercia de Newton para calcular el desplazamiento del misil en tres dimensiones y situarlo en la debida trayectoria hacia su objetivo.

La precisión de tales armas puede verse afectada por errores de muy diverso origen. Errores que proceden del sistema de guía inerte y errores que se producen en su retorno a la atmósfera. La desviación del objetivo por cualquiera de estas causas es de 100 metros o más en algunos de los ICBM de la generación actual. Los errores en el sistema de guía se derivan de imperfecciones en los giroscopios o en los acelerómetros y pueden ser constantes, acumulativos con el tiempo o aleatorios; se producen por vibraciones, golpes o cambios de la aceleración. Los errores que hacen referencia a la reentrada en la atmósfera tienen su origen, a su vez, en cambios atmosféricos imprevisibles sobre el objetivo (vientos y variaciones de la densidad atmosférica) y en incertidumbres resultantes de la ablación del cono de proa del RV. Puesto que los RV del ICBM entran en la atmósfera a una velocidad de 7000 metros por segundo, el cono de proa se consume al atravesar el RV las capas de aire. Su ablación suele producirse de una manera bastante imprevisible y asimétrica, generando inestabilidades aerodinámicas que alteran la precisión del RV. Además, ciertos factores climatológicos de importancia –como fuertes lluvias o nieve– pueden aumentar considerablemente la ablación, con la consiguiente pérdida de precisión del RV.

Entre otras causas de menor importancia se pueden incluir errores en el modelo de gravedad empleado por el sistema de guía de inercia, errores en la fusión de la cabeza nuclear y errores en la determinación del objetivo. Cada grupo de éstos agrega varias decenas de metros al "presupuesto de error" del conjunto del sistema. La determinación de la posición original y de la velocidad



1. PRIMEROS EFECTOS de la explosión atmosférica, a baja altura, de una carga nuclear de 550 kilotones, secuenciados en esta doble página. Se presupone que la dimensión del ingenio explosivo y la altura de la explosión son las adecuadas para un ataque de ICBM sobre silos de Minuteman norteamericanos. Una milésima de segundo después de la explosión de la ojiva, la temperatura de la bola de fuego que se forma oscila entorno a los 400.000 grados Celsius; la sobrepresión (el aumento de la presión por encima de la presión ambiente) se cifra en 7000 kilogramos por centímetro cuadrado. Hasta una distancia de 800 metros del punto de la explosión se expande radiación (sobre todo neutrones y rayos gamma) capaz de destruir una ojiva nuclear. A las 50 milésimas de segundo, el radio alcanzado por la bola de fuego en su rápida expansión ronda ya los 500 metros y la temperatura en su interior ha descendido a 75.000 grados Celsius. La sobrepresión en el frente de la onda de choque (que en este estadio coincide con la superficie de la bola de fuego) es de 42 kilogramos por centímetro cuadrado y, en su perímetro, el viento se expande con una fuerza de varios miles de kilómetros por hora. Cuando ha transcurrido un segundo, la bola de fuego tiene un radio de 900 metros, una temperatura interna de 10.000 grados Celsius y una temperatura de superficie de 6000 grados. El frente de choque se expande ahora más rápidamente que la bola de fuego y ha alcanzado ya un radio de 1400 metros, distancia en la que la sobre-

del lanzador es de extrema importancia en sistemas de lanzamiento desde plataformas móviles, como es el caso de los SLBM, instalados en submarinos, pero es despreciable en el de los ICBM, instalados en silos. A esos posibles fallos se debe la poca confianza puesta en la precisión de los actuales SLBM en su actuación sobre silos "reforzados" de misiles.

El montante de fuentes de error entre un misil balístico y otro es aleatorio, de modo que si se disparase hacia un solo blanco una cantidad considerable de tales armas, éstas tenderían a caer dispersas al azar en torno al blanco. La amplitud de esta dispersión se mide por un parámetro conocido por

UN MINUTO

CINCO KILOMETROS

10 SEGUNDOS

FRENTE DE CHOQUE



presión es de 2,8 kilogramos por centímetro cuadrado. A los 10 segundos, la bola de fuego ha llegado a su radio máximo, un kilómetro, y comienza a ascender. La temperatura de superficie de la bola es de unos 2000 grados Celsius y, el radio del frente de choque, de unos cinco kilómetros. (Sólo se alcanza a ver en el extremo superior de este cuadro.) La sobrepresión en el frente de choque es de 0,35 kilogramos por centímetro cuadrado. Vientos verticales de una velocidad aproximada de 600 kilómetros por hora comienzan a levantar del suelo polvo y otros escombros hacia el tallo de la nube ascendente. Al cabo de un minuto, la nube ha crecido, con su característica forma de hongo, hasta un radio de 2,5 kilómetros; su centro alcanza los 6,5 kilómetros de altura. La bola de fuego no produce ya radiaciones a longitudes de onda visibles. Vientos verticales de varios centenares de kilómetros continúan levantando grandes partículas hacia la nube y su tallo. Cualquier RV (*reentry vehicle*, la porción que se desprende del misil y lleva la carga nuclear) que entrase en su inmediata proximidad en estos primeros estadios tendría muchas posibilidades de ser destruido por radiaciones nucleares o térmicas, vientos o colisiones con las mayores partículas levantadas por la explosión de una ojiva de la primera oleada. Incluso en posteriores estadios, partículas más pequeñas levantadas por la ojiva de la primera oleada podrían afectar a las de la segunda. Tales efectos destructivos se han denominado "fratricidas".

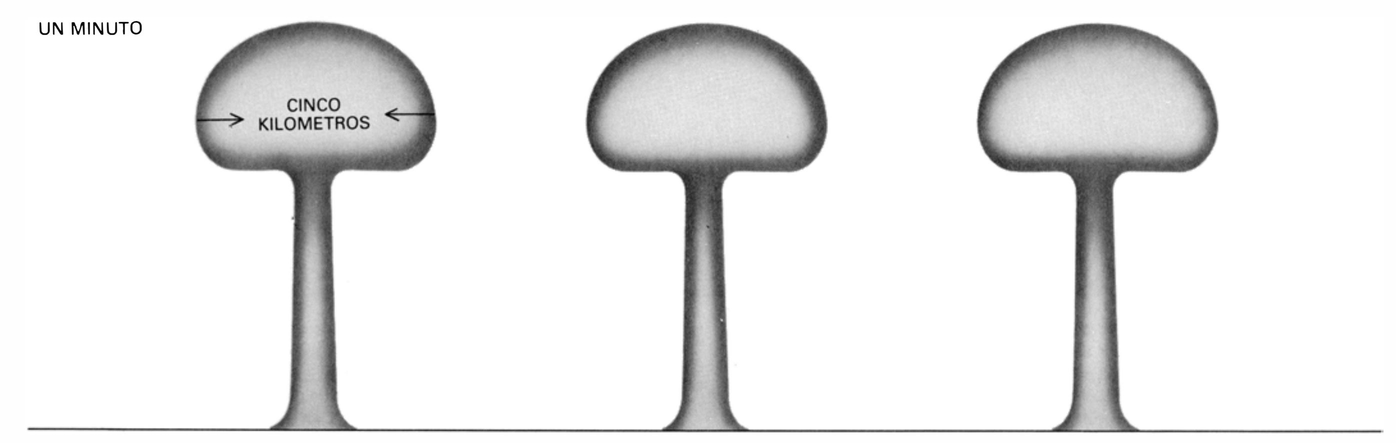
probable error circular (PEC): el radio de un círculo cuyo centro es el punto medio de impacto dentro del cual caerían el cincuenta por ciento de los RV. El PEC mide, pues, la precisión con que un misil descarga su explosivo. Con las evaluaciones del PEC, de la capacidad explosiva y de la seguridad del funcionamiento de una determinada arma se puede calcular la probabilidad que tiene ese arma de destruir un objetivo de una determinada solidez. A título de ejemplo, se considera que el misil soviético SS-19 Mod 3 [modificado] posee un PEC de 250 metros y cada una de sus ojivas una potencia de aproximadamente 550 kilotones. El SS-18 Mod 4 goza de una precisión similar y carga ojivas de potencia también similar, pero lleva diez ojivas para objetivos independientes. Suponiendo un perfecto funcionamiento de su mecanismo, las ojivas de este tipo tendrían el 63 por ciento de probabilidad de destruir un silo de Minuteman "reforzado" para soportar hasta 140 kilogramos por centímetro cuadrado. Suponiendo ahora una independencia estadística, dos armas de este tipo tendrían un 85 por ciento de probabilidad de destruir el mismo silo. Un cálculo así indicaría que con 2000 ojivas la Unión Soviética destruiría el 90 por ciento de los 1000 Minuteman norteamericanos in situ. Este tipo de resultados teóricos alarmistas ha conocido una amplia difusión.

Pero no parece muy razonable esperar que un sistema tan complejo como un moderno ICBM tenga una perfección de funcionamiento en su mecanismo de casi el 100 por ciento. Un cálculo más fiable la sitúa en el 75 por ciento para los ICBM soviéticos, lo que reduciría al 72 por ciento la probabilidad de destrucción de sus ojivas, cifra que se ajusta al 70 o 75 por ciento que hoy señala la Junta de Jefes de Estado Mayor de los Estados Unidos.

Más aún, esa cifra no recoge el amplio abanico de incertidumbres que hay que tomar en consideración al hacer un cálculo de los resultados tras un ataque contra silos. Quien detenta la responsabilidad de la planificación de un ataque así tiene siempre que enfrentarse además con otras incertidumbres: cuál sea la precisión real, la perfección del funcionamiento y la potencia de sus armas en condiciones operativas y también con qué solidez el adversario protege sus silos. Otros factores menos conocidos, como las interferencias "fratricidas" entre las armas empleadas en un ataque, influirán de una manera poderosa en el resultado. Como mostraremos, estas incertidumbres harán totalmente imposible que, con la capacidad actual de sus ICBM, los dirigentes soviéticos lleguen a tener una confianza razonable en destruir más de la mitad de los misiles norteamericanos instalados en tierra.

Volvamos en primer lugar a la principal causa de incertidumbre: la precisión de las armas atacantes. El volumen de errores de un sistema ICBM no tendrá correlación estadística de un misil a otro; a ello contribuye el azar de la caída descrita por el PEC. Pero, como ocurre en otros sistemas electromecánicos complejos, los ICBM no sólo presentan errores fruto del azar, sino también errores sistemáticos. Por consiguiente, aunque los cálculos habituales sobre vulnerabilidad del ICBM hacen presumir que el centro del impacto estará directamente sobre el objetivo, la caída real se producirá a menudo un tanto fuera del mismo. Se llama sesgo o desviación a la distancia entre el objetivo y el punto medio del impacto. Si la situación que se ha calculado para un determinado objetivo está a 20 metros al sur del punto real, es probable que también se calcule mal la ubicación de otros objetivos en la misma zona: silos instalados en el mismo campo de misiles, por ejemplo. De modo parecido, los errores en materia de gravedad de los ICBM lanzados desde silos situados muy próximos entre sí en territorio del atacante guardarán una estrecha correlación, por causa de errores debidos a los vientos existentes y a las variaciones de la densidad de la atmósfera con que se encontrarán al volver a entrar en ésta en la misma zona del territorio atacado. Por el contrario, con pruebas y análisis en tiempos de paz, cabría eliminar los más importantes sesgos y desviaciones del propio sistema de guía.

Cualquier desviación de menos de un centenar de metros no afecta sustancialmente al éxito del ataque contra silos "reforzados" para resistir hasta 140 kilogramos por centímetro cuadrado. La razón de ello reside en que cualquier arma que cayese a una distancia del objetivo tres veces superior a la antedicha lo destruiría. Sin embargo, por la propia naturaleza de la desviación, y atendiendo a la ley de los grandes números, no será fácil poner un límite a las magnitudes de las desviaciones que se producirán en un ataque real. Si un error entre dos ojivas es mero fruto del azar, un ataque que comprenda 2000 ojivas producirá 2000 valores independientes de esa variable; como ocurriría si se tirase 2000 veces un dado; son muy escasas las probabilidades de una desviación importante del resultado medio. Por otra parte, si un error guar-



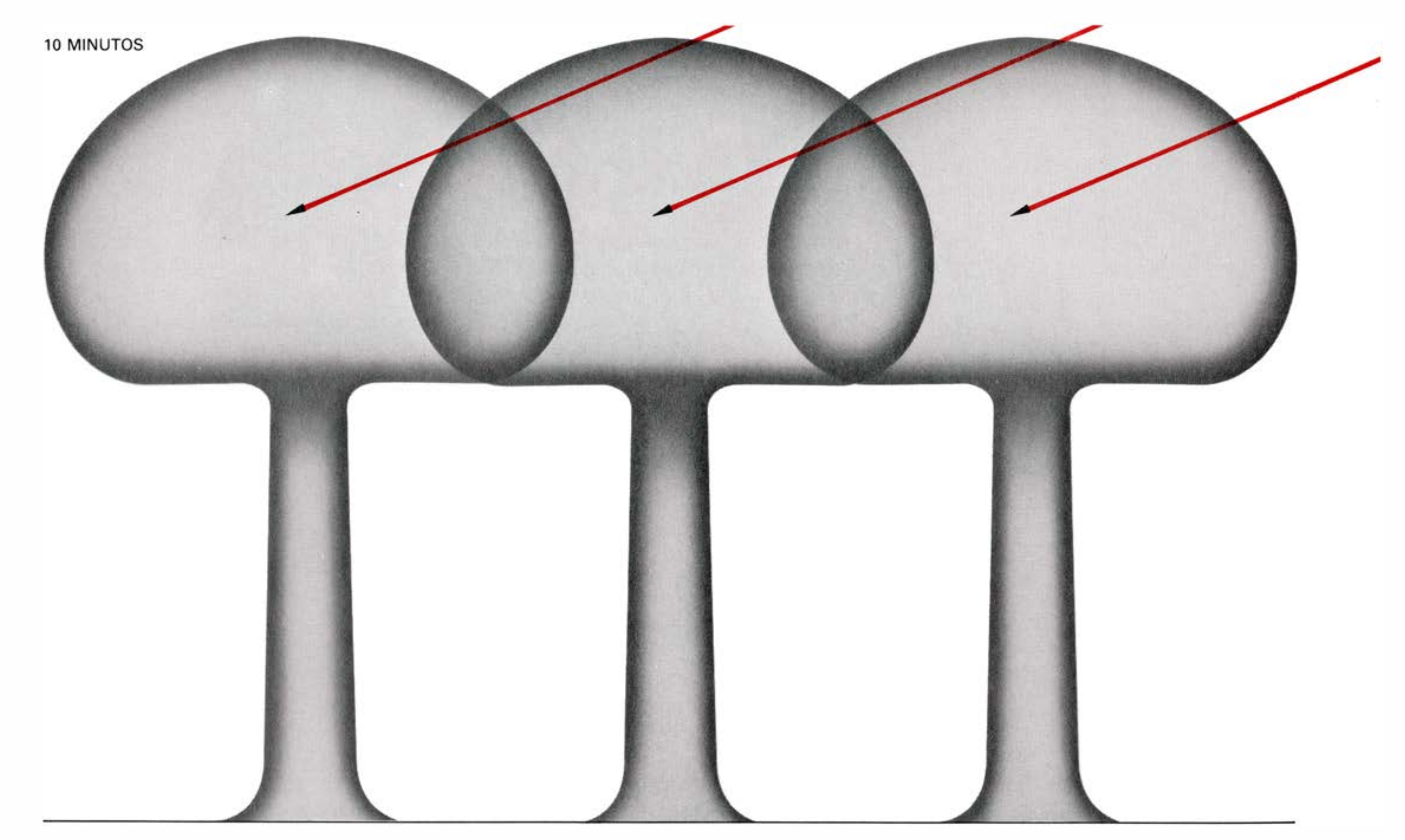
2. NUBES EN FORMA DE HONGO que ascienden tras las explosiones simultáneas, a baja altura, de varias cabezas nucleares de 550 kilotones. Aparecen aquí en dos estadios: un minuto (*izquierda*) y diez minutos después de la explosión (*derecha*). Las explosiones se provocaron a un intervalo de ocho kilómetros, que es la distancia media entre dos silos de misiles en un campo de Minuteman. Pasados diez minutos, las nubes alcanzan una envergadura de unos ocho kilómetros y su parte más alta se ha estabilizado a la altura de unos 18. La ilustración muestra varios RV que se dirigen hacia objetivos más dis-

da una correlación absoluta en alguna porción de la fuerza, para esa porción el ataque sólo da lugar a un valor de la variable. Igual que de una tirada de dos dados pueden salir un dos o un doce en lugar del valor más probable del siete, así, desde un punto de vista realista, no pueden descartarse en el ataque de un solo misil desviaciones de bastante más de 100 metros.

Hay también alguna incertidumbre en los cálculos del PEC para todo ICBM. Se dispone de una amplia variedad de fuentes de información sobre la precisión de una determinada arma. Por ejemplo, las pruebas no destructoras en superficie ofrecen minuciosos detalles sobre el funcionamiento de cada componente del sistema de guía, que pueden combinarse para componer una evaluación aproximativa de la precisión de conjunto del sistema. La realidad sigue siendo que muchas e importantes causas de error –como los fallos atribuibles al retorno a la atmósfera– no se pueden comprobar de forma fehaciente en superficie. Además, la interacción entre los diversos componentes del sistema de guía en el violento entorno de vibración, golpes y aceleración del lanzamiento de un cohete es en extremo compleja. De ahí que sólo una cantidad estadísticamente importante de pruebas realistas en vuelo de sistemas completos aporten evaluaciones precisas del PEC de un misil balístico.

Por lo general, estas pruebas se dividen en dos grandes categorías: (1) pruebas de investigación y desarrollo, que sirven para sugerir cambios en el diseño y para proporcionar una primera información sobre qué precisión puede alcanzar el misil, y (2) pruebas operativas, que proporcionan evaluaciones de la precisión y del funcionamiento de la fuerza desplegada. Puesto que un misil balístico suele costar más de 10 millones de dólares, los programas de vuelos de pruebas se ven generalmente limitados por barreras presupuestarias. Por ello, los Estados Unidos y la Unión Soviética tienden a efectuar un número relativamente pequeño de vuelos de pruebas de cada sistema de misiles. Por ejemplo, en los primeros estadios de su despliegue, lo normal es que el ICBM norteamericano se someta a 25 o 30 ensayos, que proporcionan una primera evaluación de la precisión y del funcionamiento operativo del sistema. A estas pruebas siguen entre 5 y 10 ensayos operativos cada año del ciclo vital del sistema, para controlar todos los cambios que resulten de un prolongado almacenamiento.

Los vuelos de prueba operativos, en Estados Unidos, se dirigen a asegurar que los ensayos sean lo más realistas posible. Los ICBM a ensayar se escogen al azar entre la fuerza operativa. Aquellos en los que ha recaído la elección se ponen entonces alerta en un silo y se preparan para su inmediato lanzamiento. Con ello se trata de mantener en forma la dotación humana y la electrónica del silo de lanzamiento. Si el misil no se pone en estado de alerta, se considera que ha fallado y no se prosigue su prueba. Si supera esa fase, se traslada desde su silo operativo al centro de pruebas de la base aérea de Vandenberg, en California. Los RV se envían a unas instalaciones especiales, donde se les retiran sus ojivas termonucleares para volvérselas a incorporar con equipo de telemetría. Entonces,



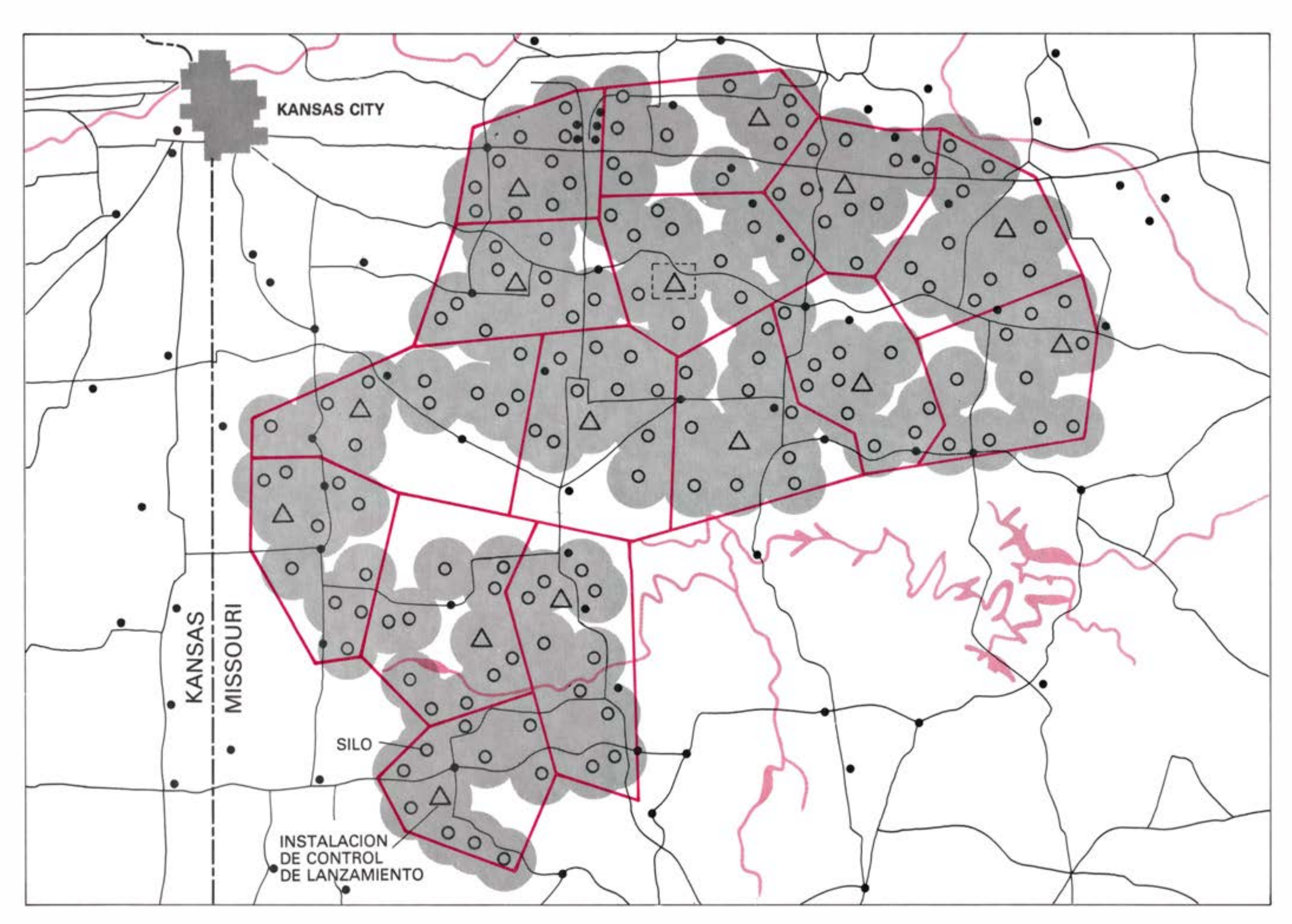
tantes. La colisión a alta velocidad con pequeñas partículas de las nubes podría desencadenar efectos fratricidas catastróficos sobre los RV de la segunda oleada, alterando seriamente su precisión, si es que no los destruyen. Los ICBM que superaron el ataque de la primera oleada podrían lanzarse sin peligro a través de la capa de nubes antes de que las ojivas que se aproximan en segunda oleada entraran, también sin peligro, en la zona de sus objetivos. El atacante sólo podría arrasar entonces silos vacíos. (Contra esta táctica, el atacante puede detonar cargas nucleares que impidan la salida de los ICBM.)

una dotación seleccionada al azar entre las secciones de misiles operativos lanza el misil con sus RV modificados desde el silo de pruebas de Vandenberg y los RV entran a la atmósfera sobre el atolón de Kwajalein, en el Pacífico.

El equipo de telemetría a bordo del misil proporciona por radio una detallada información sobre determinados parámetros: flujo de carburante, avance del cohete, comportamiento de los componentes de su guía y vibraciones y golpes a que el misil se ve sujeto. Esta información es recogida por estaciones norteamericanas de control de pruebas y por barcos espías soviéticos situados en el Pacífico a tal objeto. Además, el vuelo se controla cuidadosamente desde radares y telescopios ópticos emplazados en Vandenberg y en Kwajalein. La información se somete a un profundo análisis a lo largo de un período de varios meses. Resultado de todo ello es que la información generada por el vuelo de prueba va mucho más allá de saber si el misil funcionó o no y de a qué distancia aterrizó del objetivo previsto.

Sin embargo, persisten importantes incertidumbres. En primer lugar, el número de vuelos de prueba de todo el sistema es aún bastante reducido. En segundo lugar, y a pesar de los esfuerzos por lograr la mayor verosimilitud posible, estas pruebas en tiempos de paz difieren aún de forma apreciable de las operaciones en tiempos de guerra. Por ejemplo, los RV norteamericanos de pruebas vuelven a entrar en la atmósfera sobre el atolón de Kwajalein, una zona donde las condiciones atmosféricas se cuentan entre las más tranquilas del mundo. Además, dado que el campo de gravedad de la tierra varía de un lugar a otro, las pruebas hechas sobre una única trayectoria no pueden por sí solas proporcionar evaluaciones de posibles errores gravitatorios sobre otras trayectorias.

Lo que es aún más significativo, la distancia sobre la que se ensayan la mayoría de los misiles soviéticos es de 6500 kilómetros, mientras que en tiempos de guerra muchas trayectorias rozarían los 10.000. Aunque los soviéticos realizan varios vuelos de pruebas en todo su alcance para evaluar sus sistemas ICBM, la mayoría de los ensayos se limitan a distancias cortas. Esta diferencia de alcance tiene un notable efecto sobre la práctica totalidad de las causas de error del sistema. Aunque puedan hacerse ajustes sobre las evaluaciones de la precisión que de aquéllas resultan, a partir de un minucioso modelo matemático del comportamiento del sistema, inevitablemente aparecerán otras incertidumbres. Y no puede dejar de tenerse en cuenta que, en un ataque a gran escala contra silos, el PEC podría ser un 10 por ciento mayor –es decir, peor– que el PEC calculado a partir de disparos efectuados sobre distancias de prueba. Para las ojivas a que



3. EXTENSION DEL MANTO DE NUBES resultante de las explosiones de la primera oleada de ojivas nucleares en un hipotético ataque contra fuerzas en un campo de Minuteman próximo a la base área de Whiteman (Missouri). Unos diez minutos después de la explosión de una ojiva nuclear de 550 kilotones sobre cada uno de los 150 silos de Minuteman y las 15 instalaciones de control de lanzamiento en el área del objetivo, las nubes en forma de hongo habrían cubierto prácticamente todo el campo. Si se hiciese explosionar las ojivas de la primera oleada a nivel del suelo, el manto de nubes resultante contendría millones de toneladas de polvo y escombros. Incluso a la máxima altura posible para un ataque sobre tales objetivos "reforzados" con ojivas de 550 kilotones, esa manta de partículas contendría aún cientos de miles de toneladas de material succionado por las bolas de fuego en su ascensión.

nos hemos estado refiriendo, una variación negativa del 10 por ciento reduciría por sí sola la probabilidad de destrucción de doble golpe contra un silo de Minuteman del 72 al 66 por ciento.

Las incertidumbres sobre el funcionamiento de un ICBM son independientes de las incertidumbres relativas a su precisión. Las simulaciones en tierra son considerablemente más útiles para evaluar los porcentajes de fallo que para evaluar la precisión del sistema de arma. Por ejemplo, a los sistemas de guía del Minuteman norteamericano se les somete a pruebas periódicas en vuelos simulados, en los que el sistema de guía se ve sujeto a vibraciones y golpes semejantes a los del vuelo de un misil. Pero, como ocurre en todo sistema técnico sensible y complejo, sólo a través de ensayos de todo el sistema se puede obtener un alto grado de confianza en las evaluaciones de la totalidad de su funcionamiento. Además, las evaluaciones sobre el funcionamiento global de los ICBM tienen que tomar en consideración un amplio espectro de valores humanos que se verían implicados en el ataque: cualquier golpe contra silos a gran escala exigiría la cooperación de cientos de personas, cuyo comportamiento resulta imprevisible en semejantes circunstancias. Así pues, ese 10 por ciento probablemente constituye una evaluación conservadora de la incertidumbre global del funcionamiento. Una vez más, una variación desfavorable del 10 por ciento en el funcionamiento que se presume para un SS-19 Mod 3 reduciría el número de silos Minuteman destruidos en un ataque "dos sobre uno" del 72 al 67 por ciento.

Las evaluaciones del poder explosivo de las ojivas termonucleares contienen a su vez sus propias incertidumbres. Son éstas de dos tipos interconexos: en primer lugar, la incertidumbre sobre los efectos concretos de ojivas de una determinada potencia; en segundo lugar, la incertidumbre sobre las evaluaciones de la potencia de una determinada clase de ojivas. La realización de ensayos de explosivos nucleares en tiempo de paz se ve limitada por razones de costos, seguridad, impacto sobre el medio ambiente, instrumentación y política. Además, está también limitada por tratados, como el de Prohibición Limitada de Pruebas Nucleares de 1963 y el aún no ratificado Tratado Umbral de Prohibición de Ensayos.

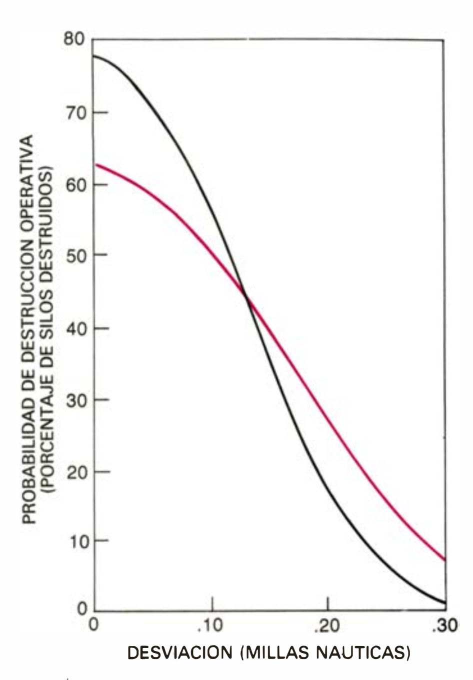
La medición de las sobrepresiones en el alcance extremo necesario para destruir un moderno silo de ICBM se ha visto limitada por problemas instrumentales y ausencia de un estado de necesidad. A principios de los años 1960, cuando norteamericanos y soviéticos llevaron a cabo las últimas explosiones nucleares en la atmósfera, los objetivos de interés más duros eran más "blandos" que los actuales silos de misiles. Por eso, hay pocas posibilidades de obtener datos de pruebas nucleares para sobrepresiones de más de 14 kilogramos por centímetro cuadrado y aún menos para las de más de 35 kilogramos por centímetro cuadrado. Los datos de pruebas nucleares existentes para esas altas sobrepresiones muestran un altísimo grado de dispersión y a menudo no concuerdan con las previsiones teóricas. Aunque pueden obtenerse algunos datos adicionales de pruebas con explosivos químicos, surgen grandes incertidumbres al extrapolar tales datos a las previsiones de los efectos de las armas nucleares del orden del megatón. La Agencia de Inteligencia de la Defensa de Estados Unidos ha declarado que las evaluaciones norteamericanas de las sobrepresiones que cabe esperar de las explosiones nucleares a ciertos niveles tienen una incertidumbre del 20 por ciento en más o en menos.

La segunda incertidumbre al evaluar el poder explosivo de las ojivas termonucleares, relativa a la potencia media de una determinada clase de ojivas, tiene un origen parecido: el número de pruebas de cualquier arma nuclear es pequeño y existen importantes dificultades instrumentales para medir la energía producida por una explosión nuclear subterránea de ensayo.

La potencia que se espera del diseño de una determinada arma puede evaluarse de modo teórico, pero este tipo de análisis tiene sus trampas, como se vio en la ojiva desarrollada para el RV Mark 12, recientemente desplegado en los misiles norteamericanos Minuteman III. Las tres primeras pruebas de este arma mostraron que su potencia era bastante inferior a la prevista, y hubo que modificar el diseño original hasta que, en un cuarto ensayo, el arma alcanzó toda su potencia. Aunque éste pueda ser un caso extremo, la incertidumbre sobre la potencia media de cualquier clase de ojivas será del 10 por ciento, por lo menos.

En aras de la simplicidad, combinaremos los dos tipos de incertidumbre en el poder explosivo de una ojiva, describiéndolos como variaciones en potencia efectiva. Dado que la sobrepresión máxima tiene que ver con la potencia explosiva, una variación del 20 por ciento en la sobrepresión produciría una variación del 20 por ciento de la potencia explosiva; una estimación conservadora de la incertidumbre total de la potencia efectiva de una determinada arma podría ser entonces del 25 por ciento. Una variación desfavorable del 25 por ciento en la potencia de las armas implicadas reduciría la eficacia del hipotético ataque que hemos descrito del 72 al 66 por ciento.

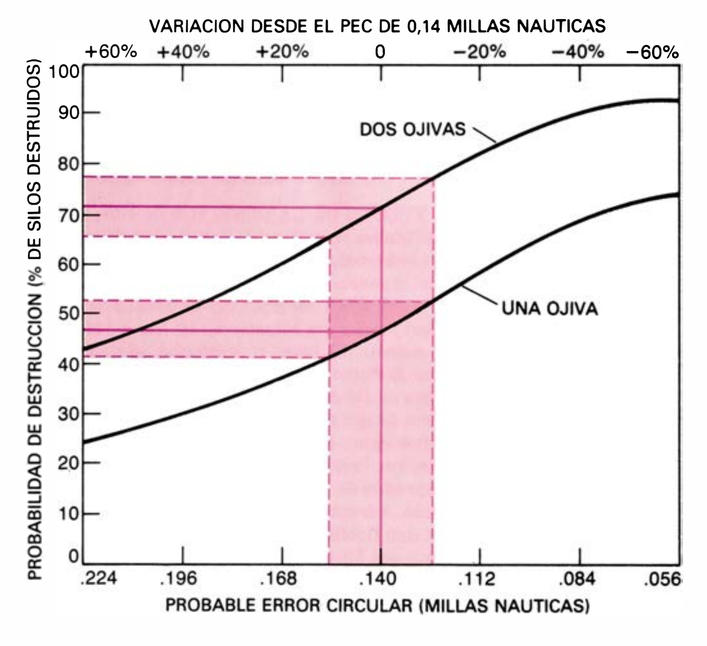
Aunque la precisión, el funcionamiento y la potencia de los ICBM no son seguros, un país que contemple la posibilidad de un ataque se ve obligado a ensayarlos en tiempos de paz. Lo que no rige para la solidez de los silos que

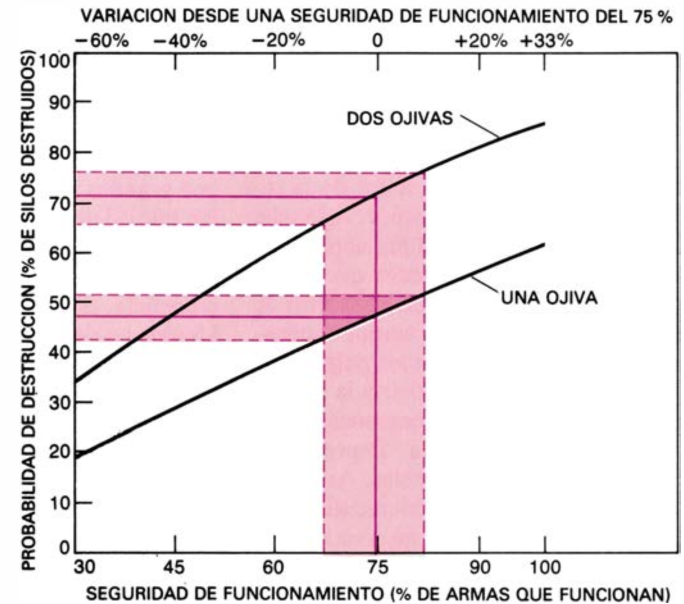


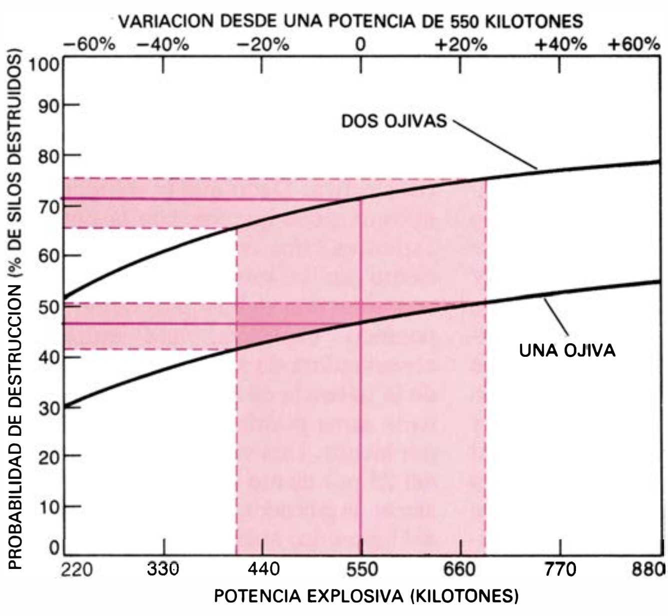
4. EFECTO DE LA DESVIACION sobre las probabilidades operativas de destrucción de una ojiva que se aproximara en un hipotético ataque contra silos. Aparece en este gráfico con referencia a dos RV: una ojiva soviética SS-19 Mod 3 (*curva de color*) y una ojiva Minuteman III Mark 12A (*curva en negro*). (El sesgo o desviación se define aquí como la distancia que hay desde el objetivo hasta el punto medio de impacto de una muestra al azar de ojivas dirigidas a ese objetivo.) Se supone que en ambos casos las ojivas han sido lanzadas sobre objetivos "reforzados" para sobrevivir a una sobrepresión de 140 kilogramos por centímetro cuadrado. Además, se supone que los sistemas operan con una fiabilidad del 100 por ciento, que las dos ojivas del SS-19 tienen una potencia explosiva de 550 kilotones y, 350, las de los Minuteman III y que el probable error circular es de 0,14 millas náuticas para el SS-19 y de 0,10 para el Minuteman III. (El probable error circular –PEC– se define como el radio de un círculo dentro del cual caerán la mitad de las ojivas dirigidas al objetivo.) Es de señalar que para ambos misiles una desviación de media milla náutica (unos 100 metros) apenas incidiría en la probabilidad de destrucción; una desviación mayor sí tendría consecuencias.

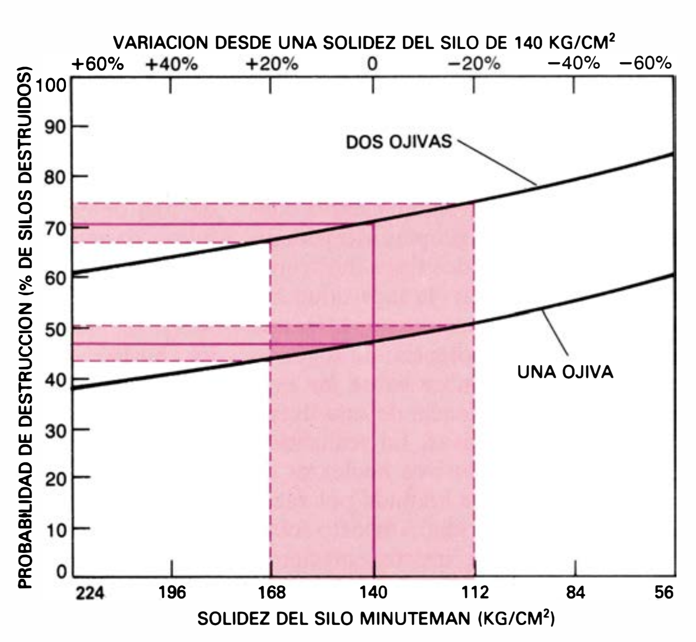
se van a atacar. La sobrepresión a la que un silo sucumbirá depende, en primer lugar, de características técnicas de la puerta de hormigón reforzado situada en lo alto del mismo, y es en extremo difícil obtener información fiable de sus características. En eso puede estribar la mayor de las incertidumbres, desde el punto de vista del atacante. Los propios Estados Unidos admiten que la solidez de sus silos no goza de todas las garantías y seguridad ante imponderables. Ningún silo se ha visto aún expuesto a prueba nuclear. Las evaluaciones de la solidez de los silos se basan totalmente en consideraciones teóricas sobre su estructura y pruebas de modelos a escala con explosivos químicos. Así pues, desde el punto de vista de quien ataque, es posible que la incertidumbre de la solidez de los silos a destruir sea al menos del 20 por ciento, si no bastante mayor. Si los silos objeto del ataque resistieran sobrepresiones de un 20 por ciento superiores a lo esperado, se reduciría, del 72 al 68 por ciento, la eficacia del ataque hipotético que hemos descrito.

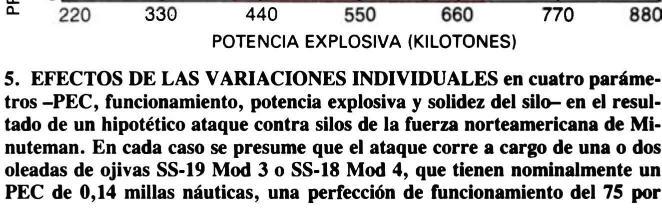
Otra importante incertidumbre surge de la interferencia entre los centenares de ojivas que intervendrían en un ataque. Hasta aquí hemos supuesto que las ojivas empleadas en un ataque contra silos serían estadísticamente independientes entre sí; en otras palabras, que la explosión de una ojiva no incidiría en las demás participantes. Pero de ningún modo ocurriría así: las explosiones termonucleares pueden tener efectos enormemente destructivos sobre otros RV, lo que constituye un efecto denominado irónicamente "fratricida".










5. EFECTOS DE LAS VARIACIONES INDIVIDUALES en cuatro parámetros –PEC, funcionamiento, potencia explosiva y solidez del silo– en el resultado de un hipotético ataque contra silos de la fuerza norteamericana de Minuteman. En cada caso se presume que el ataque corre a cargo de una o dos oleadas de ojivas SS-19 Mod 3 o SS-18 Mod 4, que tienen nominalmente un PEC de 0,14 millas náuticas, una perfección de funcionamiento del 75 por ciento y una potencia de 550 kilotones. A los silos de los Minuteman se les ha asignado una solidez nominal de 140 kilogramos por centímetro cuadrado. Las franjas de color suave señalan cuán importantes son los efectos sobre las probabilidades de destrucción de una incertidumbre evaluada conservadoramente en un 10 por ciento para el PEC y para el funcionamiento, del 25 por ciento para la potencia explosiva y del 20 por ciento para la solidez del silo.

En sus primeras milésimas de segundo, una explosión termonuclear libera un intenso estallido de radiaciones: neutrones, rayos X y rayos gamma; las radiaciones crean, a su vez, un poderoso impulso electromagnético. A esta breve explosión de radiación sigue la rápida expansión de una bola de fuego de gases calientes y comprimidos. La bola de fuego se expande a una velocidad superior a la del sonido, alcanzando un radio de centenares de metros en menos de un segundo. Puesto que los gases supercalentados de su interior son menos densos en el aire que la rodea, la bola de fuego empieza a ascender rápidamente, cual si de un globo de aire caliente se tratara. Ello desencadena un viento vertical de varios centenares de kilómetros por hora; de hecho, basta el efecto del viento provocado por una explosión de medio megatón para levantar una piedra de dos toneladas de peso. En el caso de que la explosión de las armas se produzca a una altura óptima para el ataque a un silo, endurecido para soportar 140 kilogramos por centímetro cuadrado, la bola de fuego estará en contacto con la tierra varios segundos; consiguientemente, los poderosos vientos ascendentes absorberán hacia arriba millares de toneladas de polvo y escombros, creando la famosa nube en forma de hongo.

La nube asciende a gran velocidad. En un minuto alcanza una altura de varios kilómetros. Entonces, frena su ascensión, llegando a su máxima altura a los diez minutos de la explosión. En el caso de un arma de medio megatón, la parte superior de la nube se estabilizará a los 18 kilómetros de altura, con su base inferior unos ocho más abajo. En esos diez minutos, la nube habrá cubierto una extensión de unos 100 kilómetros cuadrados. La nube de tal explosión será tan grande que, en caso de ataque a silos de ICBM en un campo estadounidense de Minuteman, las nubes de las explosiones sobre los distintos silos crearían un manto de polvo sobre todo el campo.

La importancia del polvo y de los escombros producidos por la explosión se deriva del hecho de que los RV atraviesan la atmósfera a gran velocidad. Cuando, al entrar en la atmósfera en una zona en la que ya se ha producido una explosión, el RV se topa con la nube, estará desplazándose a unos 6 kilómetros por segundo. Por consiguiente, si el RV colisionase con una partícula que pesase varios gramos, lo probable es que se destruyera inmediatamente, pues semejante colisión se produciría a una velocidad varias veces superior a la de una bala de fusil. Las partículas más pequeñas y el polvo de la nube podrían tener efectos abrasivos catastróficos sobre el RV: el efecto de ese paso a gran velocidad a través de la nube equivaldría a la exposición a un chorro de arena extraordinariamente intenso. La ablación resultante del cono de proa del RV reduciría gravemente su precisión y, en casos extremos, podría suponer la combustión del RV. Además, aquella primera explosión (o explosiones) habría modificado por completo, y de forma imprevisible, la densidad y los perfiles del viento del área, lo que tendría efectos decisivos sobre la precisión de los RV.

Como es lógico, el problema del "fratricidio" tendría que considerarse cuidadosamente al planificar la temporalización de cualquier ataque que exigiese más de un RV dirigido a cada silo. Si se dirigiesen al mismo tiempo dos a cada silo y no fallase ninguno de ellos, la radiación de la primera ojiva que estallase destruiría la segunda. Si el ataque se organizase en dos oleadas separadas por varios segundos, los RV de la segunda se encontrarían con las bolas de fuego, fuertes vientos y partículas letales causadas por las explosiones de la primera. Aunque las altas temperaturas y los vientos de escape generados por la explosión se disiparían a los pocos segundos, las partículas suficientemente grandes para afectar fatalmente a los RV tardarían 20 minutos en caer a tierra y las nubes de polvo y partículas más pequeñas se mantendrían hasta dispersarse por los vientos de la atmósfera. Por tanto, salvo que el atacante dejase transcurrir al menos varios minutos entre la primera y la segunda oleada, para que las partículas mayores hubiesen caído de la nube antes de la llegada de la segunda oleada, parece poco probable que una gran proporción de esta segunda alcanzase su objetivo y lo hiciese con la precisión necesaria para destruirlo. Aunque el atacante dejase transcurrir varias decenas de minutos entre las dos oleadas, la segunda tendría aún que enfrentarse con el manto de polvo y con graves perturbaciones atmosféricas, producto de la explosión, o explosiones, de la primera oleada.

Al no haberse sometido a prueba todavía, no podemos evaluar con exactitud los efectos del fratricidio sobre los RV; hay que conformarse con estimaciones de mera aproximación. Supongamos, en el caso de una segunda oleada que entrase en la atmósfera 10 minutos después de la primera, que su paso por la nube de polvo y los choques con grandes partículas destruyesen sólo el cinco por ciento de los RV y que, por término medio, el efecto de las alteraciones atmosféricas duplicase la fracción del PEC atribuible a la entrada en la atmósfera. Se trata, a nuestro parecer, de una evaluación bastante conservadora del fratricidio que esa oleada sufriría, dadas las condiciones extremas que hemos descrito. Estos efectos reducirían por sí solos la probabilidad de destrucción de un ataque en dos oleadas llevado a cabo por ojivas SS-19 Mod 3 sobre los silos de Minuteman del 72 al 65 por ciento.

El fratricidio introduce una incertidumbre aún más significativa: si el atacante tiene que dejar que transcurran varios minutos entre la primera y segunda oleada de su golpe, es muy posible que los ICBM que sobrevivan a la primera hayan partido ya de sus silos antes de la llegada de la segunda oleada. A su salida del silo, un ICBM se desplaza mucho más lentamente que un RV que retorna a la atmósfera; por tanto, las partículas y el polvo de la nube tienen comparativamente menos efecto sobre él. Los ICBM que sobrevivan podrían ser lanzados sin peligro antes de que la segunda oleada de RV entrara, también sin peligro, en la atmósfera. Aunque una explosión nuclear próxima que no haya conseguido destruir el silo puede impedir que el ICBM despegue inmediatamente, en esos casos en los cuales la ojiva de la primera oleada no tuvo éxito el ICBM continuará sin daño alguno en su silo reforzado. En el escenario de ataque aquí considerado, y que incluye armas de una fiabilidad en su funcionamiento del 75 por ciento, ello significaría que, entre la primera y la segunda oleada, podría lanzarse un mínimo del 25 por ciento de los ICBM. En combinación con los efectos fratricidas anteriormente postulados, reduciría la porción de silos de ICBM destruidos en el hipotético ataque al 56 por ciento. Sin embargo, puesto que la vulnerabilidad de un ICBM aumenta rápidamente en cuanto abandona su silo protector, el atacante podría impedir el disparo de esas armas detonando sobre el campo de silos más ojivas, a intervalos regulares, para destruir cualquier ICBM lanzado desde su silo protector. A esta táctica se la denomina, en el lenguaje militar, "atenazar contra el suelo" (*pindow*).

Hemos pasado revista, una por una, a las incertidumbres que rodean un ataque contra silos. Pero en cualquier ataque real contra silos todas las incertidumbres se darían a la vez, dificultando así las previsiones de qué podría ocurrir tras el ataque. Imaginemos que todas las variables reseñadas se tornasen desfavorables para el atacante; aun no existiendo una desviación considerable sólo se destruiría el 45 por ciento de los silos de Minuteman norteamericanos. Por poco probable que nos parezca que tal ataque estuviese sujeto a fuertes cambios desfavorables en todos los parámetros al mismo tiempo, habría que señalar que una variación desfavorable de dos parámetros cualquiera, sumada al fratricidio, rebajaría la eficacia del ataque a menos del 55 por ciento, aun cuando se ignorasen las posibilidades de error de desviación o las que tiene el ICBM de emprender la huida entre la primera y la segunda oleadas. Somos conscientes de que los grados de incertidumbre que hemos postulado pecan de conservadores. Así pues, la sección de estado mayor responsable de la planificación difícilmente pondría una esperanza razonable en que un ataque en dos oleadas con ojivas SS-19 Mod 3 y SS-18 Mod 4 destruyese más de la mitad de la fuerza de misiles Minuteman.

Puede que el responsable de la planificación optase por lanzar más de dos ojivas contra cada silo, o por lanzar ojivas mayores. Sin embargo, en el caso de ataque contra la fuerza norteamericana de Minuteman, ninguna de esas opciones se presentaría como muy atractiva. Por ejemplo, una tercera oleada se encontraría con efectos fratricidas procedentes de las dos primeras y daría mayor margen de tiempo, después de la segunda, para lanzar los ICBM supervivientes. Aunque los misiles no pudiesen huir, una tercera oleada sólo incrementaría el porcentaje global de silos destruidos del 45 al 57 por ciento. Dado que los 1000 Minuteman norteamericanos albergan 2100 ojivas, ello significaría un gasto de 1000 cabezas soviéticas más para destruir 250 ojivas norteamericanas, de tal modo que un ataque semejante desarmaría antes a la Unión Soviética que a Estados Unidos. Aparte de sus ICBM armados con MIRV, la Unión Soviética tiene 100 SS-19 y SS-18 dotados con una sola ojiva de gran potencia. Demasiado pocos para un ataque contra los 1000 silos de ICBM norteamericanos.

Los dos argumentos anteriores no valen para ciertos objetivos, pocos aunque importantes. Si Estados Unidos desplegase 100 ICBM MX en silos de Minuteman, tal como ha propuesto la Administración Reagan, podrían emplearse para atacarlos los 100 misiles soviéticos equipados con grandes ojivas. Si se dotase a esas armas con los precisos sistemas de guía que despliegan esos mismos misiles en su versión MIRV, el ataque de una sola ojiva de ese tipo sobre cada silo de MX, seguido del de otra ojiva de medio megatón, destruiría el 77 por ciento de las fuerzas MX, aunque se diesen condiciones adversas y fratricidas. Puesto que cada MX lleva 10 ojivas, ello significa que, incluso en condiciones no óptimas, 200 ojivas soviéticas podrían destruir 770 ojivas MX. El empleo, por parte de la Unión Soviética, de otras ojivas adicionales proporcionaría unas posibilidades de destrucción aún mayores.

HIPOTESIS	DESVIACION (MILLAS NAUTICAS)			
	0	.05	.10	.15
FUNCIONAMIENTO AL 100 POR CIENTO	86%	84%	76%	63%
FUNCIONAMIENTO AL 75 POR CIENTO	72%	70%	62%	50%
LIGERO FRATRICIDIO	65%	62%	56%	45%
VARIACIONES DESFAVORABLES	45%	43%	38%	31%

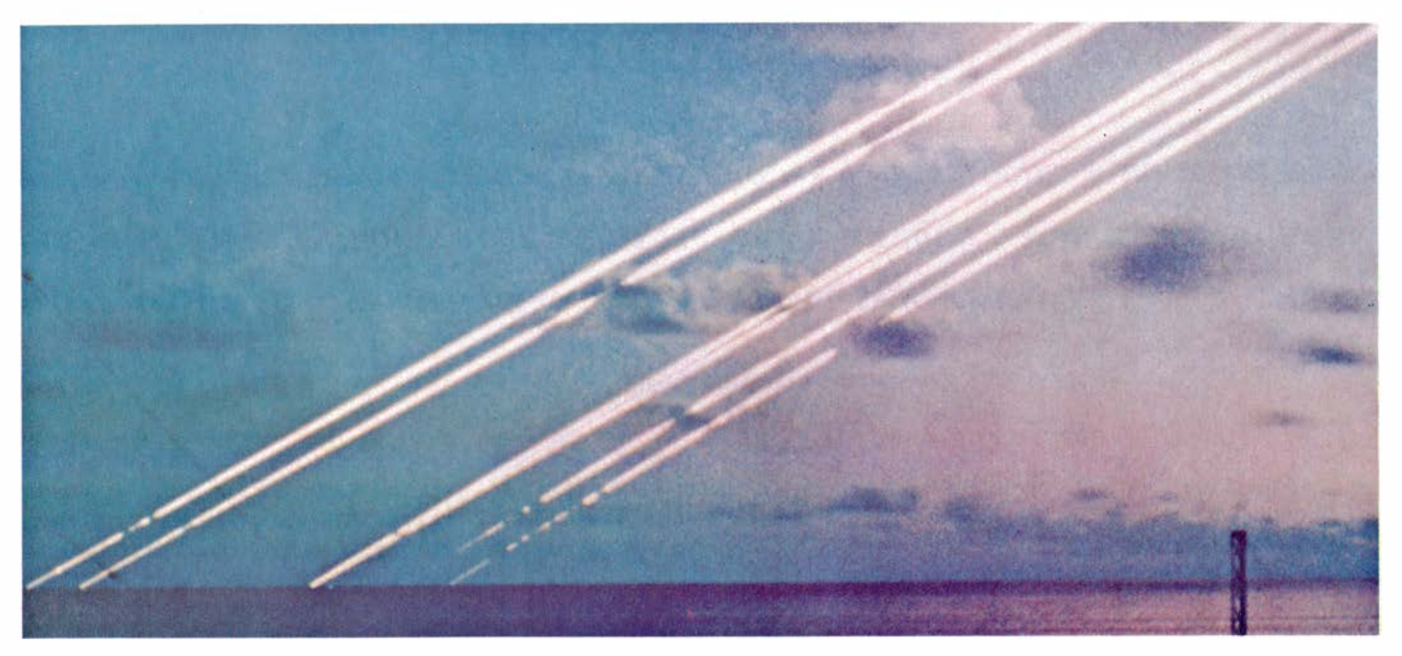
6. EFECTOS DE LA CONJUNCION DE INCERTIDUMBRES en el resultado de un ataque contra silos. Las columnas corresponden a cuatro proyecciones de desviación diferentes. Las filas indican los efectos de las hipótesis propuestas acerca de las capacidades técnicas de las armas empleadas en el ataque (y, en un caso, de la solidez de los silos que se atacan). La primera línea muestra el resultado en términos de probabilidades de destrucción operativa de un ataque ideal, con armas cuyo funcionamiento es totalmente seguro, sobre silos sin variación en su solidez con respecto al valor nominal. La segunda muestra los resultados de un ataque algo más realista, con armas de una seguridad nominal en su funcionamiento del 75 por ciento. La tercera fila informa del efecto producido por una mínima cantidad de fratricidio. La cuarta fila muestra el resultado de un ataque en el que un ligero fratricidio se combina con variaciones desfavorables (para el atacante) en todos los parámetros restantes. Aquí, la precisión y el funcionamiento resultan un 10 por ciento por debajo de los valores nominales y los silos a atacar un 25 más sólidos de lo previsto. Sin esa desviación significativa, no se destruiría más del 45 % de los silos de Minuteman.

Así pues, aunque la actual fuerza ICBM soviética representa una seria amenaza para un pequeño número de objetivos de especial valor, ello no le daría, a quien planificase un ataque, una confianza razonable de destruir de modo significativo más de la mitad de la actual fuerza ICBM norteamericana. La práctica, tan extendida, de referirse al resultado probable con dos importantes cifras, sin mencionar la incertidumbre que acecha, constituye una falsificación grotesca: desde un punto de vista realista, todo lo que se puede decir es que el resultado de tal ataque sería *probablemente* la destrucción de entre un 50 y un 90 por ciento de los ICBM norteamericanos. De lo que se infiere que se ha exagerado sin tasa la magnitud de la amenaza de los ICBM soviéticos.

Pero esta conclusión, que pudiera aliviar alguna conciencia, quizá no tenga una validez duradera. La tecnología de las armas estratégicas progresa por días y los adelantos previsibles en los sistemas para situar las armas sobre sus objetivos pueden alterar drásticamente la situación descrita. En el pasado, la precisión de los ICBM, tanto americanos como soviéticos, se multiplicó por dos aproximadamente cada siete años. El refinamiento técnico de las armas estratégicas podría avanzar menos deprisa, a medida que se estrecha el campo del progreso, pero no hay razón para esperar que en los próximos años el modelo difiera mucho.

Puede llegar el día, a finales de la presente década o en la próxima, en que la Unión Soviética haya desplegado una fuerza ICBM dos veces más precisa que la actual. Además, en Estados Unidos –y posiblemente también en Rusia– se están desarrollando RV diseñados para penetrar en las nubes de polvo. Estos cambios tecnológicos podrían menguar de modo significativo el efecto de las incertidumbres que se han descrito. Si el PEC del SS-19 Mod 3 se redujese, por ejemplo, a la mitad, el ataque de dos a uno acabaría con más del 80 por ciento de la fuerza Minuteman estadounidense, incluso dándose cierto fratricidio y variaciones adversas en los parámetros del golpe. En síntesis, una incertidumbre del 10 al 20 por ciento no modificará las cosas si el arma es dos veces más precisa de lo necesario para destruir su objetivo.

Pero las incertidumbres puramente técnicas que se han descrito constituyen la punta del iceberg. Quien proyecte un ataque contra silos tendrá



7. IMPACTO DE UN ATAQUE CONTRAFUERZAS, recogido en esa fotografía de la Fuerza Aérea de los Estados Unidos, en la que se observa seis MIRV MARK 12 cayendo, sin carga real, sobre sus objetivos en el atolón de Kwajalein, en el Pacífico occidental. Los RV iban alojados en dos misiles balísticos intercontinentales (ICBM) y fueron lanzados, durante unas maniobras, desde la base aérea de Vandenberg (California). Su trayectoria de entrada en la atmósfera es típica de un ataque contra silos en el que, desde diferentes misiles, se dirigen dos ojivas hacia cada objetivo "reforzado".

también que considerar incertidumbres de otro tipo, centradas sobre todo en la reacción del país agredido. Una nación que almacene ICBM tiene la opción de adoptar la política de "lanzamiento por alerta" o "lanzamiento ante ataque": lanzar los misiles inmediatamente de recibir la información de que otro país ha disparado sus misiles contra él. Estados Unidos ha renunciado a esa política, ante la posibilidad de un error catastrófico, pero el estratega difícilmente podría llegar a confiar en que esa política valga en caso de ataque. Habrían de transcurrir casi 30 minutos entre la primera detección de un lanzamiento masivo y la explosión de las ojivas sobre los silos. Es imposible predecir qué harían los responsables políticos durante ese período y en los siguientes.

Si además de los ICBM instalados en tierra se tienen en cuenta las restantes fuerzas nucleares estratégicas estadounidenses, a la sección de estado mayor responsable se le planteará un serio problema. En primer lugar, los miles de armas nucleares instaladas en submarinos en navegación serán invulnerables a un ataque, al menos en un futuro previsible. En segundo lugar, resultaría imposible destruir conjuntamente la fuerza ICBM y los bombarderos estratégicos estadounidenses. Una parte de la fuerza bombardera está en alerta permanente de 15 minutos, lo que significa que los únicos misiles que podrían destruirla antes de emprender el vuelo serían misiles lanzados desde submarinos que cubriesen la distancia en un rápido vuelo. Sin embargo, los actuales SLBM soviéticos son muy poco seguros para ser eficaces contra silos reforzados, lo que significa que habría que atacar a los silos con vuelos de mayor duración. Si esto ocurriese, las armas nucleares empezarían a estallar sobre las bases de los bombarderos norteamericanos más de 15 minutos antes de que los RV dirigidos sobre los silos pudiesen alcanzar sus objetivos, lo que permitiría el lanzamiento de los ICBM antes de la llegada de los RV.

Aceptemos que se venzan esas dificultades: un ataque que supusiese más de 2000 explosiones nucleares de una potencia media de un megatón, a ras del suelo, causaría entre 20 y 40 millones de bajas civiles. Tamaña agresión no puede valorarse de forma realista como una "acción de cirugía nuclear". No cabe imaginar que Occidente no respondiese, apoyándose en una parte de los miles de ojivas que aún le quedaran en bombarderos y submarinos. Con toda probabilidad, el conflicto experimentaría una rápida escalada hasta un intercambio estratégico con empleo de todas las fuerzas, lo que destruiría al atacante tanto como a la víctima del golpe inicial.

Con los avances de la tecnología, y en un futuro más lejano, podrían reducirse esas incertidumbres implícitas. El desarrollo de SLBM más precisos, que ya está en marcha en Estados Unidos, debilitará la relación entre la fuerza de alerta de bombarderos y la fuerza de ICBM. Si se desplegaran RV extremadamente precisos en su maniobra, los silos podrían destruirse con armas de mucha menor potencia, reduciendo en gran medida el número de bajas civiles que podría causar el ataque. Aunque ello puede parecer deseable, a primera vista, la posibilidad de una acción auténticamente "quirúrgica", combinada con una mayor confianza en las previsiones de sus resultados, aumentaría la tentación de lanzar semejante ataque, lo que rebajaría el umbral nuclear y aumentaría las probabilidades de guerra.

Así pues, aunque la situación actual es más estable de lo que suele creerse, los progresos de la tecnología armamentista no ofrecen unas perspectivas tranquilizadoras. Cabe la posibilidad de que unas rigurosas limitaciones de las pruebas y del despliegue de misiles balísticos pongan coto a muchos de esos pocos deseables avances. Lejos de encerrar a Estados Unidos en una situación de vulnerabilidad, tales limitaciones, de tener efecto, podrían evitar un rápido deterioro de la seguridad estadounidense, que en caso contrario se producirá inevitablemente en los próximos años. Las limitaciones de ensayos y el despliegue de misiles balísticos podrían constituir un importante componente de los esfuerzos para el control de armamentos y merecen un estudio más cuidadoso que los que hasta el momento se han realizado.

Las salmueras

Constituyen un recurso natural del que se ha obtenido productos básicos de la industria alimentaria y química. Tras una etapa de recesión, el descubrimiento de nuevos tipos y componentes en ellas abre un amplio abanico de posibilidades

Salvador Ordóñez

Los kukakuka, conjunto de tribus caníbales del este de Nueva Guinea, famosos por su habilidad en el manejo de las armas, son unos especialistas consumados en la fabricación de sal a partir de salmueras artificiales. Su régimen alimentario está formado exclusivamente por tubérculos y presenta un fuerte déficit en sodio. Las distancias relativamente grandes al mar y la situación casi permanente de guerras tribales hacen que la sal marina no pueda comercializarse hacia el interior. Los kukakuka fabrican la sal que necesitan, y que constituye la base de su sistema monetario, a partir del lixiviado de las cenizas de la planta *Coix gigantea* Koening esc Rob. La transplantan a zonas regables y, una vez cortada, seca y calcinada, se deposita sobre recipientes de calabaza haciendo colar agua a través de las cenizas. El agua salada así obtenida se transporta en cañas de bambú al taller del "especialista", que dispone de un horno de evaporación, donde durante cinco días, la salmuera, vertida en unos moldes a temperaturas inferiores a ebullición, acabará dando un precipitado salino compacto. Estas barras constituyen la base de los intercambios mercantiles entre las tribus.

La sal común, producida a partir de salmueras marinas, era conocida en China 3000 años a. C. La Biblia recoge numerosas referencias a la sal; así, según el *Génesis*, la mujer de Lot es convertida en estatua de sal por ser un alma incrédula; la tradición popular llegó a asociar un cono de sal, existente en el suroeste del mar Muerto, con el testimonio de aquel castigo divino. Es posible que, aunque no haya testimonios escritos, el hombre haya conocido la sal desde muy antiguo. Quizá la descubrió de una manera casual en la costa de una laguna salada; atraído sin duda por las eflorescencias que se forman en la parte emergida y sorprendido por su sabor, comprobaría cómo los animales muertos accidentalmente se conservaban en su seno. Sea como fuere, la consideró dotada de poderes mágicos y la utilizó en todo tipo de ceremonias religiosas.

En el antiguo Egipto la industria del vidrio, que luego desarrollarían los fenicios, se basaba en las sales carbonatadas sódicas que se precipitaban en algunos lagos efímeros, situados unos 100 kilómetros al noroeste de El Cairo, alimentados periódicamente por el río Nilo.

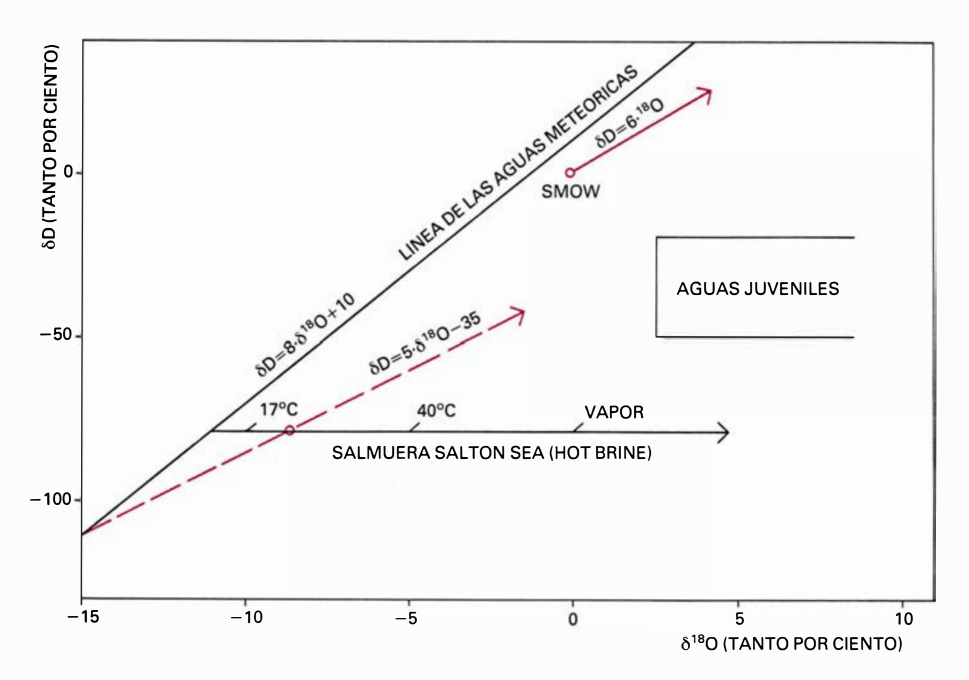
La producción semiindustrial de sal empezó utilizando zonas naturales comunicadas episódica o restringidamente con el mar, donde, debido al calor solar, se concentraban las salmueras y se depositaban las sales durante la estación seca. Poco a poco ese sistema natural se mejoró levantando estanques artificiales con el objeto de aumentar la producción y mejorar la calidad del producto. En los países de clima húmedo se desarrolló la técnica de precipitación de sales, que se hacía entonces por ebullición. Así, alrededor de los centros productores de carbón al norte del Reino Unido, y ante un bloqueo del mercado de sales por parte de los países productores (Francia y España), se favoreció durante los siglos XVI y XVII toda una industria basada en la utilización del carbón como combustible y del agua de mar como salmuera. El descubrimiento de yacimientos de sal gema, a finales del siglo XVII, provocó la desaparición de esta técnica.

En el Código de las Siete Partidas se dice: "La renta de las salinas son de los emperadores y reyes, a quienes se otorgan, para que con su producto se mantuviesen honradamente, pudiendo ampliar sus tierras y reinos, y guerrear contra los enemigos de la fe, y evitar imponer a sus pueblos nuevos tributos y gravámenes". En España la fabricación y distribución de la sal común estuvo sujeta a monopolio real durante siglos, con pequeñas etapas liberalizadoras. Los manantiales salinos se conocían aquí desde antiguo: salinas de Atienza y salinas de Espartinas y Almoguera, cuya explotación venía regulada por una Real Provisión de Carlos I en el año 1555. Fue en la "laguna" de Espartinas donde fue descubierta la thenardita por Roda en 1850. Las salinas de Torrevieja y La Mata, las mayores sin duda de Europa, administradas por el Ministerio de Hacienda directamente producen del orden de la mitad de la sal marina de España. (Hasta 1759 se intentó convertirlas en albufera; desde entonces se renunció a dicho proyecto y en 1841 la Empresa Arrendataria de la Renta de la Sal, del Marqués de Salamanca, reformó el método de explotación, recolectando la sal desde barcas en lugar de sacarla a lomos de caballerías. La producción media anual, que para los años 1953-1970 era de 400.000 toneladas, aumentó espectacularmente con la adición de una salmuera artificial obtenida por disolución del diapiro salino de "el Pinoso", situado a 50 kilómetros de Torrevieja. El tratamiento de las salmueras concentradas permite obtener, además de sal común, varios subproductos: bromo, ácido bromhídrico, bromuro sódico, potásico y amónico, sulfato magnésico y cloruro magnésico, en cantidades suficientes para el consumo interior español.)

A comienzos de nuestro siglo, se registraba en Estados Unidos una demanda importante de fertilizantes, tanto nitratos como sales potásicas, estas últimas importadas de Alemania.

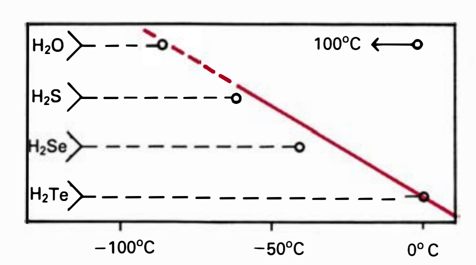
1. LOS MANANTIALES SALINOS se forman por la acción de las aguas infiltracionales sobre materiales que contienen sales en forma diseminada o masiva. Estos manantiales se han venido explotando para la obtención de sales. La fotografía de la derecha muestra una explotación artesanal lisboeta, en la que se aprovechan las aguas procedentes de un diapiro salino de facies Keuper para la obtención de sal común.

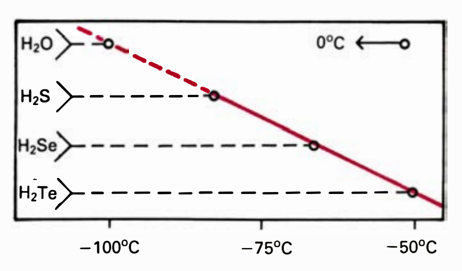




2. "MEMORIA" DE LAS AGUAS NATURALES. Con dicho nombre se quiere significar que su composición isotópica refleja perfectamente su origen y evolución. Las aguas marinas, volumétricamente las más importantes, tienen una composición isotópica muy constante. Todos los datos isotópicos se refieren a ellas. (SMOW abrevia "standard mean ocean water".) Las aguas meteóricas tienen una composición variable en función de la climatología. Las del mar evolucionan en su composición isotópica por evaporación ($\delta D = 6 \cdot \delta^{18}O$), al igual que cualquier tipo de aguas meteóricas ($\delta D=\delta^{18}O - 35$). Por δ D se entiende la unidad de medida de las relaciones isotópicas dada en tantos por mil: $\delta‰ = [R \text{ muestra} - R \text{ estándar}] / [R \text{ estándar}]$ y multiplicada la razón por mil. Aguas meteóricas que se infiltran en una zona de rocas silico-carbonatadas adquieren temperaturas elevadas y aumentan su contenido en el isótopo ^{18}O; puesto que las rocas contienen poca H_2O y poco hidrógeno el δD permanece inalterado.

Se estudió entonces la posibilidad de aprovechar salmueras para la obtención de sales potásicas y comenzaron a investigarse las salmueras profundas asociadas con el petróleo, así como las de algunos lagos (Searles). Hoy, a pesar de los descubrimientos de sales





3. DE LOS HIDRUROS DE LOS ELEMENTOS del grupo VI del sistema periódico es el agua el que tiene las propiedades más anómalas. Su punto de ebullición debería ser 80-90 grados centígrados bajo cero y su punto de congelación 100 grados C bajo cero. Esta anomalía se explica por la propia estructura de las moléculas del agua. Una de las propiedades singulares del agua es la de mantenerse líquida en un intervalo de temperaturas fundamental para la vida tal y como la concebimos.

potásicas de Saskatchewan (Canadá), que permiten diversificar la demanda exterior, más del 10 por ciento del consumo interior norteamericano se obtiene a partir de salmueras. Pero no sólo se explotan sales potásicas, sino que el litio, bromuros, carbonato sódico, sulfato sódico e incluso compuestos de iodo pueden recabarse en condiciones económicas de las citadas salmueras. Del lago Searles se obtiene litio a partir de salmueras con unas 70 partes por millón (ppm) de litio.

En el estado de Israel se especula en este momento con una comunicación del mar Muerto con el mar Mediterráneo, lo que determinaría que, junto a un importante complejo hidroeléctrico, aparecería la mayor planta de derivados químicos jamás concebida. Mientras estos proyectos esperan su ejecución, Israel cubre sus necesidades de sales potásicas con las salmueras del mar Muerto [véase: "El mar Muerto", por Ilana Steinhorn y Joel Gat; INVESTIGACIÓN Y CIENCIA, diciembre de 1983]. Aparte de sacar todo un conjunto de productos químicos que ofrecen una competencia insuperable para otros países y que se extraen de salmueras muy concentradas, Japón obtiene, a partir de salmueras asociadas con yacimientos de gas natural ubicados en la península de Chiba, el 55 por ciento de la producción mundial de iodo. Los nitratos chilenos, que conocieron un gran esplendor a principios de este siglo, y que los procesos industriales de fijación de nitrógeno arrumbaron, vuelven de nuevo a entrar en la escena de los recursos naturales por su alto contenido en litio (Salar de Atacama, 2000 ppm) y en iodo. En 1984 se estima una producción de 14.000 toneladas de carbonato de litio.

Pero quizás este importante pasado y presente de las salmueras sea el inicio de un futuro prometedor. Cuando en 1965 J. Swollow y J. Crease descubrían las salmueras calientes (44,8 grados C) del fondo del mar Rojo, que contenían en solución cobre, hierro, zinc, plata y plomo, en cantidades muy superiores a las del agua del mar, una nueva época nacía para el aprovechamiento de estos recursos: las salmueras como fuente de metales pesados. En la península de Cheleken (mar Caspio) se citan salmueras profundas que contienen entre 0,2-5,4 ppm de zinc y 2-77 ppm de plomo, precipitando blenda y galena, así como barita, a partir de ellas. Wolframio formando iones complejos y con concentraciones de 70 ppm se encuentra en las aguas de los lagos Owens y Searles, en los Estados Unidos. En el mar de Salton, en el suroeste de los Estados Unidos, las salmueras termales profundas contienen 1000 ppm de manganeso, 500 ppm de zinc, 20 ppm de cobre y 2 ppm de plata.

Las salmueras constituyen, pues, un importante recurso para la obtención de diferentes tipos de sales de metales alcalinos y alcalinotérreos. El alto gasto de energía invertido en la obtención de sales por procedimientos industriales ha llevado a pensar de nuevo en su explotación. Súmase a ello otro motivo de interés: contienen en algunos casos cationes metálicos, no sólo alcalinos y alcalinotérreos, sino incluso pesados; por si fuera poco, la presencia en la naturaleza de salmueras con temperaturas superiores a las ambientales nos sugiere una explotación mixta de la energía geotérmica y de las sustancias disueltas en las salmueras.

Básicamente, la salmuera es una disolución acuosa de iones y coloides, con concentraciones superiores al 35 por mil. Se trata de un límite arbitrario y meramente indicativo, pues el límite inferior de concentraciones para considerar salmuera un agua natural se refiere normalmente al agua marina estándar.

El agua, desde el punto de vista de su composición atómica, presenta un sinfín de formas, ya que tanto el hidrógeno como el oxígeno tienen varios isóto-

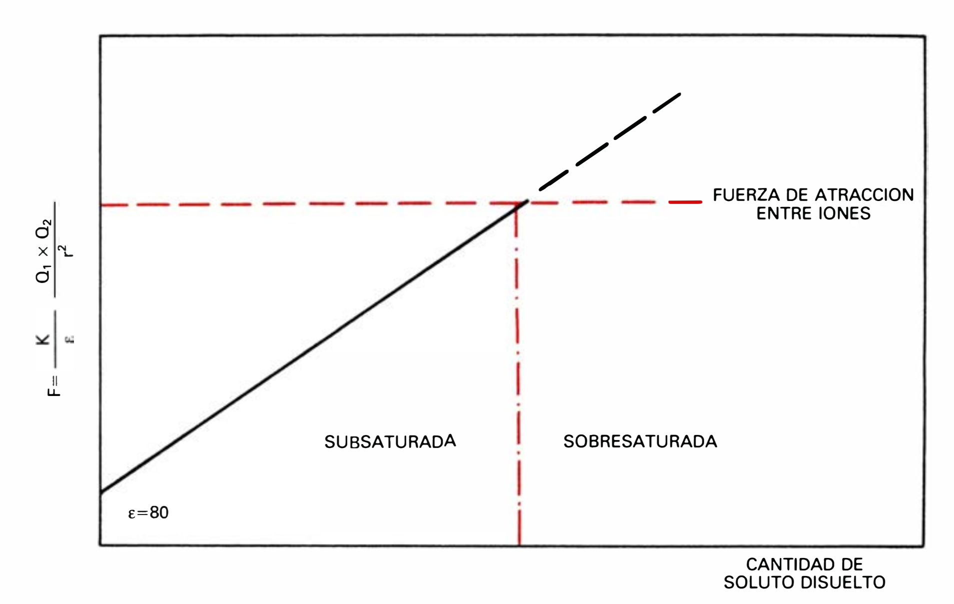
pos y todos ellos pueden formar parte de la molécula de agua. Así, el hidrógeno puede aparecer como ^{1}H (protón) y ^{2}H (deuterio). Por cada 6700 átomos ligeros (^{1}H) se presenta un átomo pesado (^{2}H). Existe además un isótopo radiactivo: ^{3}H (tritio), cuyo período de semidesintegración es de unos 12 años. El oxígeno presenta otros tres isótopos naturales estables: ^{16}O, ^{17}O y ^{18}O; aparecen en la naturaleza con abundancias medias relativas tales que, a cada 2600 átomos de ^{16}O, corresponden 10 átomos de ^{17}O y 55 átomos de ^{18}O. Se puede decir que, desde el punto de vista de la composición isotópica, no existen dos tipos de aguas iguales; así, el agua oceánica está enriquecida en isótopos pesados con respecto a las aguas continentales. El agua incorporada a los seres vivos está enriquecida en isótopos ligeros con respecto al medio ambiente en que estos seres vivos se desarrollan. Las aguas frías tienen unos valores bajos en isótopos pesados.

Las leyes que rigen el fraccionamiento isotópico son sencillas en términos generales. La cinética de las reacciones donde intervienen isótopos pesados es más lenta que sus homólogas donde intervienen isótopos ligeros. Así, en un proceso de evaporación la fracción vapor estará enriquecida en isótopo ligero y la fracción líquida quedará enriquecida en el isótopo pesado. La actividad de los seres vivos fija preferentemente isótopos ligeros por la mayor difusibilidad de éstos; lo mismo ocurre en los cristales con respecto a las soluciones en las que se generan. (Se llama difusibilidad a la aptitud de una sustancia para moverse en el seno de otra, tendiendo a homogeneizarse.)

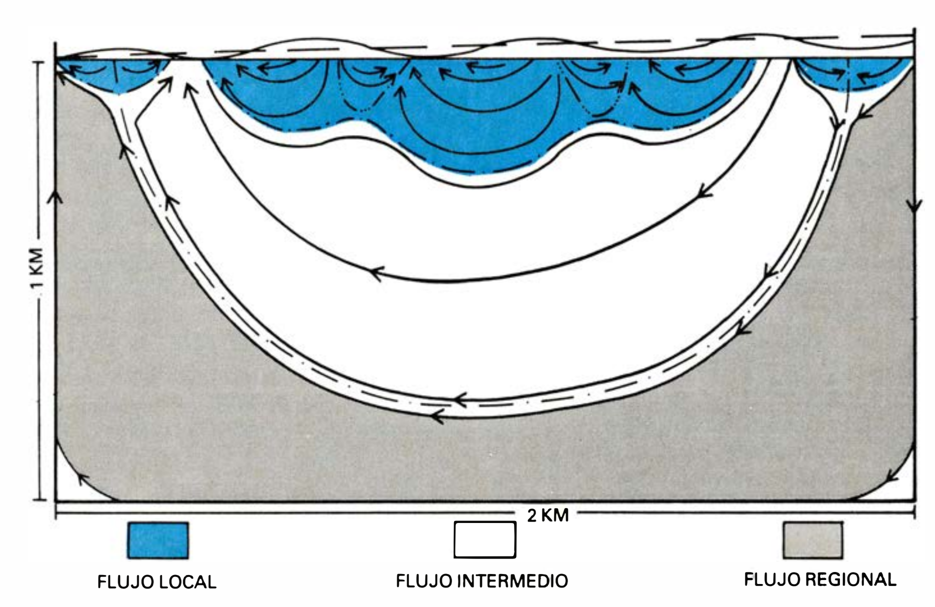
La molécula de agua tiene una estructura no lineal, con unas propiedades que la hacen anómala con respecto a otros hidruros de elementos pertenecientes al grupo VI (el del oxígeno, azufre, selenio y teluro), fundiendo e hirviendo a temperaturas muy superiores a las que teóricamente le corresponderían.

El agua pura es un dieléctrico, es decir una sustancia no conductora de electricidad y que carece, por tanto, de electrones libres, cuyos valores de la constante dieléctrica (ε) son unas 80 veces menores que el del aire. En consecuencia, las fuerzas de atracción entre las cargas (*F*), que constituye la base del enlace iónico, dominante en la estructura de muchos minerales, se debilitan unas 80 veces en el seno del agua, de acuerdo con la expresión $F = k\ (Q_1 \cdot Q_2)/\varepsilon \cdot r^2$; donde Q_1 y Q_2 son las cargas, k una constante de proporcionalidad y r simboliza la distancia entre cargas. Como resultado, se produce la disociación de los iones que formaban la molécula que presentaba enlace iónico. La cantidad de sólido disociado está autolimitada, ya que a medida que se produce la disolución el agua se vuelve conductora, o sea, disminuye su constante dieléctrica y, en definitiva, llega un momento en que las fuerzas de atracción adquieren un valor suficiente para mantener la estructura de los sólidos correspondientes.

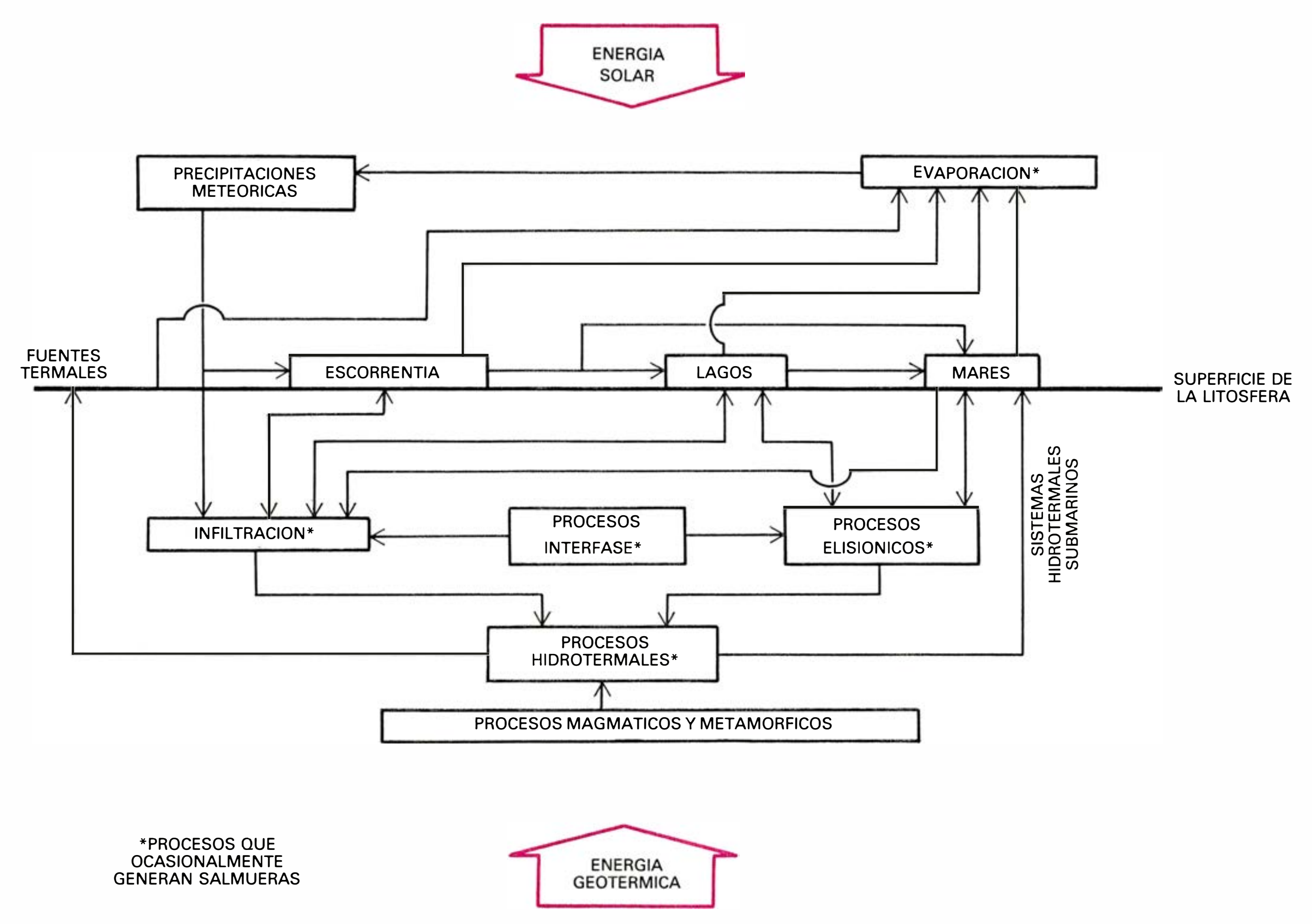
La cantidad de especie mineral disuelta depende de la naturaleza de ésta y, en definitiva, de la energía del enlace químico, siempre que el agua pemanezca en las mismas condiciones de presión y temperatura. Cuando éstas varían, la constante eléctrica varía también y por supuesto la cantidad de especie mineral disuelta. Especies minerales diferentes



4. LAS FASES CRISTALINAS SE DISUELVEN en el agua en cantidades fijas para cada temperatura y presión. El gráfico establece una relación entre la constante dieléctrica del agua y la fuerza de atracción de los iones. Como la distancia entre los iones de una estructura cristalina es prácticamente constante, y las cargas también, la única variable que puede afectar a la fuerza de atracción es la constante dieléctrica del medio. Si este medio es el agua, un dieléctrico, puede rebajar la fuerza de atracción y por tanto destruir (disolver) la estructura cristalina. A medida que los iones pasan al solvente (agua), la constante dieléctrica disminuye hasta un punto en el que se iguala la fuerza de atracción entre los iones con la fuerza necesaria para mantener la estructura cristalina, y ya no se destruye la red: la solución está saturada.



5. LINEAS DE FLUJO de las aguas infiltracionales. Se distinguen dos tipos: flujo local y flujo regional. El local es rápido, de corto recorrido y de direcciones variables en función de la topografía y de la distribución de las permeabilidades en el subsuelo, así como del sistema hidrológico y del régimen de pluviosidad. El flujo regional es lento, de largo recorrido y de direcciones muy constantes. Podemos encontrar en la zona de descarga de estas aguas de flujo local salmueras que resultan de la circulación, durante períodos de tiempo largos, atravesando las capas profundas del suelo que pueden contener rocas fácilmente solubles y son llevadas en solución a la superficie. Reelaboración de un esquema de J. Toth.



6. TODOS LOS PROCESOS que ocurren en la superficie del planeta se deben a la interacción de dos campos energéticos: la energía solar y la energía geotérmica interna del planeta. Junto con estas dos fuentes de energía, el campo gravitatorio terrestre condiciona gran parte de los procesos de circulación de las aguas superficiales e infiltracionales. En la superficie del planeta el principal proceso generador de salmueras es la evaporación, relacionada con la aridez del clima (relación entre evaporación o evapotranspiración potencial y precipitaciones). Las salmueras infiltracionales están ligadas a la disolución de rocas salinas en el subsuelo. En las aguas elisiónicas y en las termales los cambios termodinámicos son los responsables de la génesis de las salmueras.

presentan solubilidades diferentes. La expresión energética de las consideraciones anteriores viene dada por la constante de equilibrio para cada temperatura y presión, que toma el valor de 1,364 para 25 grados C y 1 atmósfera de presión. Cuando el producto de actividad iónica de los iones en solución es menor que la constante de equilibrio de la especie correspondiente se dice que la solución está subsaturada; si es igual, se dice que está justamente saturada o saturada. En el caso de que el producto de actividad iónica sea mayor que la constante de equilibrio, la solución está sobresaturada y, si se produce la nucleación adecuada en tamaño y número de núcleos cristalinos, se puede generar una especie sólida a partir de la solución.

Para precipitar una sal a partir de una salmuera podemos disminuir la cantidad de disolvente o modificar la constante de equilibrio con la presión y la temperatura. En la génesis de las salmueras naturales intervienen básicamente la evaporación, la variación de presión y la temperatura durante la evolución de las aguas naturales.

El agua se encuentra disociada en su seno, es decir, contiene en solución 10^{-7} moles de H^+ por litro y 10^{-7} moles de OH^- por litro. En definitiva, por cada 10^9 moléculas de agua dos estarán disociadas. A la concentración de H^+ (protones) en el agua se la denomina mediante la unidad convencional conocida por *p*H; ahora bien, la presencia de iones en la solución hace variar la concentración de protones en el agua y en consecuencia el *p*H: las sales de ácido fuerte tienden a aumentar la concentración de protones de la solución, mientras que las sales de ácidos débiles tienden a disminuir la concentración de protones. Posiblemente, el componente móvil más importante regulador del *p*H de los medios naturales sea el CO_2, que se puede encontrar en solución como $CO_3H^0_2$, CO_3H^- y CO^-_3. La variación del contenido en protones del medio modifica la solubilidad de las sustancias en el mismo.

Las aguas naturales disuelven el oxígeno del aire, hasta el punto de presentar una concentración mil veces menor que la existente en la atmósfera. La materia orgánica reduce por putrefacción la cantidad de oxígeno disponible en el agua. Tenemos así aguas oxidadas ricas en oxígeno y aguas reductoras pobres en oxígeno. Los iones presentes en las soluciones pueden estar en varios estados de oxidación, en función del grado de oxidación de las aguas; y, por supuesto, siempre que la energía requerida por el proceso de oxidación-reducción no alcance valores tales que produzca la descomposición electrolítica del agua. De esta manera, los iones $S^=$ en solución pueden pasar por oxidación a iones $SO^=_4$, los iones Fe^{2+} a Fe^{3+}, el Mn^{2+} a Mn^{4+}, el NH^+_4 a NO^-_3, ..., y viceversa. Se comprende, pues, que la naturaleza de los iones en solución varíen, y, por ende, las especies que se pueden obtener por saturación de los mismos en la solución.

En consecuencia, en la génesis y evolución de una salmuera, además de los parámetros de concentración, presión y temperatura, puede influir la variación de la concentración de protones y el

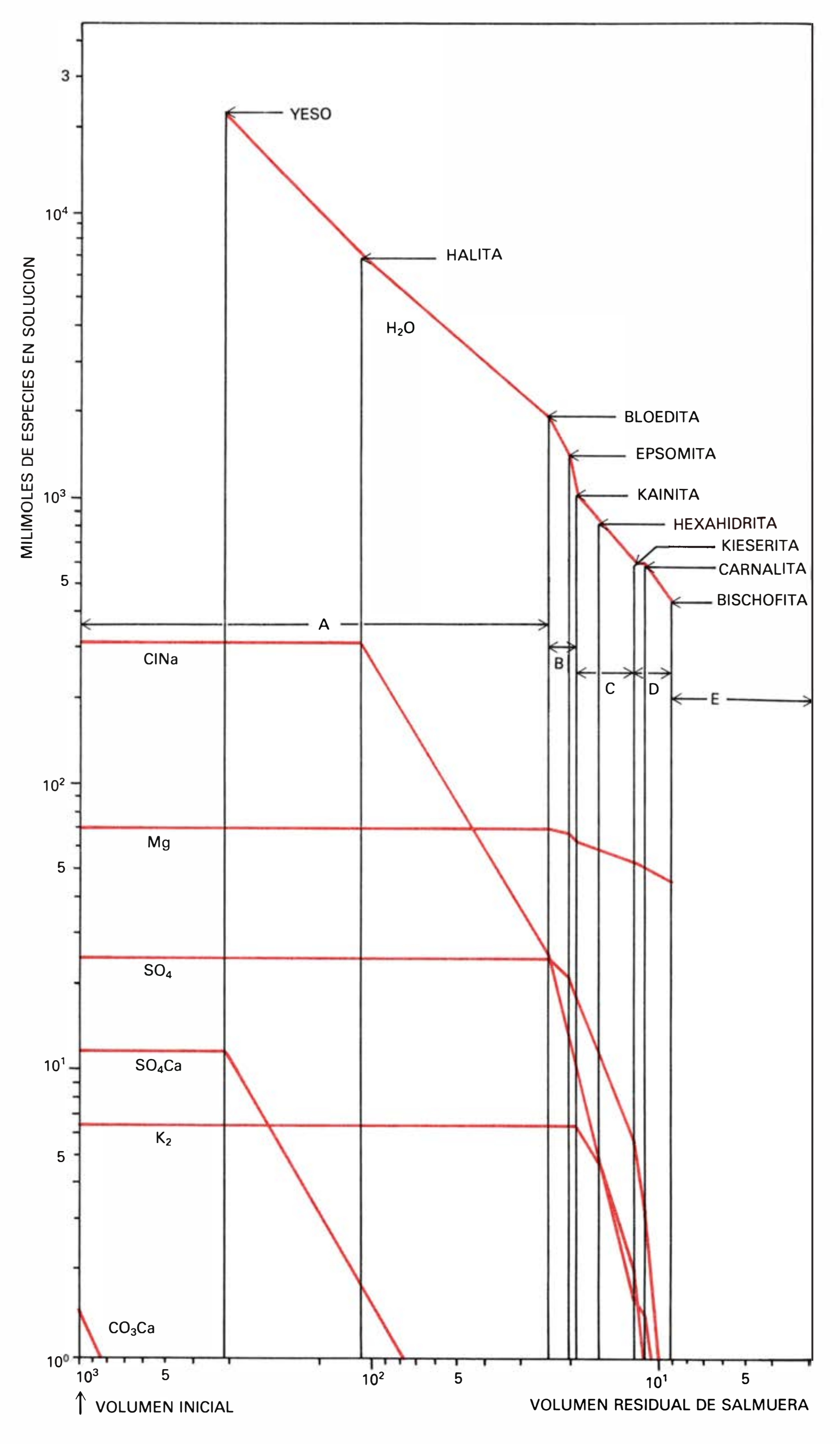
contenido en oxígeno, o si se prefiere, la cantidad de CO_2 y de oxígeno disueltos en las aguas. Quizás una de las propiedades físicas que caracterizan mejor una salmuera sea la conductividad eléctrica, (CE), expresada en mho por centímetro (unidades recíprocas de ohm). Otra propiedad importante de las salmueras es la densidad (gramos por centímetro cúbico), que aumenta regularmente con la concentración. [Aunque están en desuso, siguen utilizándose en la medida de densidades los denominados "grados Baumé" (°Bé). La relación entre la densidad dada en g/cm^3 y la dada en °Bé es la siguiente: d = 146,78 dividido por 146,78 − °Bé.]

Las aguas procedentes de las precipitaciones meteóricas, una vez que alcanzan la superficie topográfica, corren sobre ella dando lugar a las aguas de escorrentía superficial, se infiltran en el terreno o se evaporan. El que predominen unos u otros procesos depende exclusivamente de la climatología y de las características morfológicas y litológicas del sustrato sobre el que se produce la precipitación y, por supuesto, de la naturaleza de esta precipitación. Las aguas meteóricas proceden de la evaporación de las aguas de la hidrosfera (ríos, lagos, mares) e incluso de la evaporación de precipitaciones, respiración de los seres vivos, etcétera.

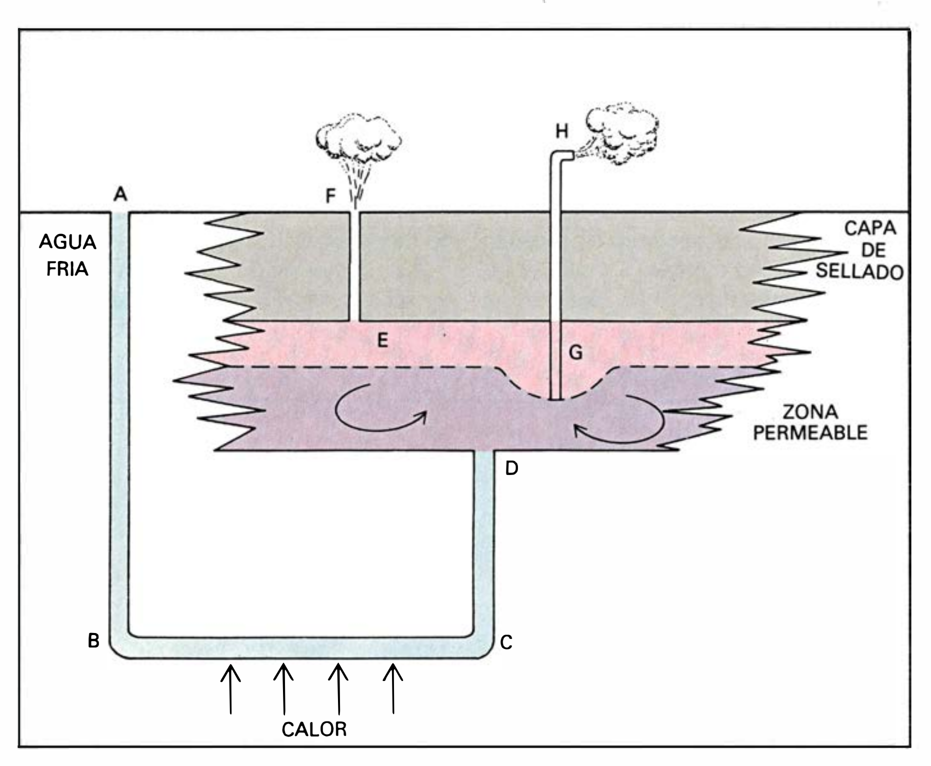
Las aguas de escorrentía evolucionan en la litosfera, yendo a parar a las cuencas lacustres o marinas. En definitiva, tenemos un ciclo externo de las aguas que está enmarcado, desde el punto de vista energético, entre la energía solar y la energía gravitacional. La variación de la energía solar es la responsable de la evaporación, mientras que la escorrentía está relacionada con la energía gravitacional; las precipitaciones participan de ambos extremos.

Existe un ciclo de las aguas subterráneas o ciclo interno que tiene como carácter distintivo no ser unidireccional; al poseer tres puntos de alimentación, provoca un número elevado de interfases. Las aguas meteóricas son oxidantes, ácidas y diluidas; al ponerse en contacto con los materiales del terreno, en función de la naturaleza de éstos, se van cargando de iones, disolviendo congruente e incongruentemente los minerales que componen las rocas. En la zona del suelo vegetal adquieren una acidez mayor, ya que la actividad microbiana genera aquí cantidades importantes de CO_2, por lo que la presión parcial de este gas aumenta extraordinariamente. Estas aguas que proceden de precipitación meteórica y que se infiltran en el terreno se denominan *aguas infiltracionales*; pueden dividirse en locales, que son las que al poco tiempo de infiltrarse afloran de nuevo en manantiales, arroyos, surgencias, etcétera, y profundas, que son aquellas que pueden llegar a tardar años antes de llegar a la superficie, siendo en cierto modo aguas fósiles.

La composición isotópica de las aguas meteóricas es relativamente

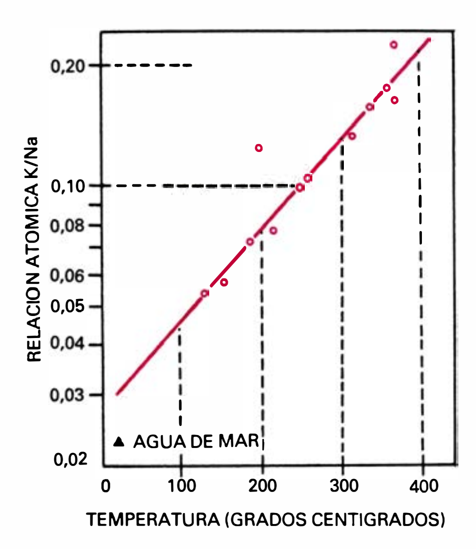


7. LA EVAPORACION del agua de mar, la salmuera universal, da lugar a una variada gama de depósitos salinos, a medida que el agua va escapando del sistema. En la parte superior se ilustra la pérdida de agua por evaporación, a partir de la cantidad inicial de un litro de agua de mar. Sobre esta línea, que no es recta, ya que el agua se pierde por evaporación, pero también como agua de cristalización, se ha señalado el comienzo de la precipitación de las diversas fases salinas, al reducirse el volumen inicial de la salmuera considerada.



8. DESARROLLO DE UN CAMPO DE ENERGIA GEOTERMICA. Las aguas meteóricas o marinas se infiltran. En el sector comprendido entre las letras *B* y *C*, se calientan por acción del flujo calorífico local; el vapor producido se almacena en un nivel permeable a partir del cual puede manifestarse de una manera natural (*F*) mediante "geyseres", o bien obtenerse artificialmente mediante sondeos. El vapor de agua o las aguas calientes pueden emplearse para la producción de energía eléctrica o calórica.

constante y las abundancias de dos isótopos radiactivos (tritio y carbono 14) permiten su datación cronológica, o lo que es lo mismo, permiten saber cuánto tiempo hace que fueron aguas meteóricas. Las aguas infiltracionales profundas acaban adquiriendo, a lo largo del tiempo de permanencia en contacto con los materiales, una salinidad elevada, sobre todo si atraviesan materiales fácilmente solubilizables. El punto de surgencia puede estar muy alejado del punto de alimentación y a veces resulta difícil establecer las líneas de flujo, a menos que se recurra a técnicas de datación de las aguas.



9. TEMPERATURA de las salmueras geotérmicas en función de la relación atómica potasio a sodio. Se expresa en una ecuación del tipo T °C = 715,89 + 197,84 ln K/Na. Esta relación empírica se explica por la variación del equilibrio entre los aluminosilicatos sódico-potásicos y el agua, a medida que aumenta la temperatura; puede constituir un magnífico geotermómetro. Algunas salmueras termales de alta entalpía (vapor) dan lugar a importantes cantidades de sales de tipo potásico (ClK).

Los sedimentos que se depositan en las cuencas sedimentarias marinas o lacustres están inicialmente formados por una gran cantidad de agua de naturaleza marina o lacustre, salada o dulce, según el medio de sedimentación. Una vez que estos sedimentos van quedando enterrados a profundidades cada vez mayores, van perdiendo la porosidad inicial, de acuerdo con una ecuación del tipo: $\Phi = A \cdot e^{-Bx}$, donde Φ designa la porosidad, A equivale a 70 en materiales detríticos finos y, a 40, en materiales detríticos gruesos; por su parte, B vale 0,001 en materiales finos y, 0,0004, en materiales gruesos. La ecuación nos pone de manifiesto que los sedimentos finos (menores de 64 micrometros, o milésimas de milímetro) durante el enterramiento se ven obligados a expulsar gran cantidad de fluidos. Estos fluidos parten de los sedimentos finos (migración primaria) y se mueven a través de sedimentos de granulometría gruesa (migración secundaria) en el sentido de mayores presiones a menores presiones y contrariamente al campo gravitatorio terrestre. Junto con los sedimentos finos, se incorpora materia orgánica (sustancias húmicas), lo que determina que las aguas que empapaban los sedimentos, inicialmente oxidantes, se vuelvan reductoras. La materia orgánica incorpora un sinnúmero de cationes pesados acomplejados en moléculas orgánicas, que en el enterramiento, y debido a la ruptura térmica de los polímeros orgánicos, quedan libres, a la vez que empiezan a aparecer radicales hidrocarburados sencillos; esta ruptura térmica está relacionada con el gradiente geotérmico (20-40 grados C por kilómetro) que es función del flujo térmico local del planeta. En definitiva, el resultado es la génesis de unas aguas reductoras ricas en metales pesados, asociadas con yacimientos de hidrocarburos; a estas aguas las denominaremos *elisiónicas*.

Las aguas elisiónicas pueden evolucionar hacia la superficie e incorporarse a las aguas de la hidrosfera. También pueden ponerse en contacto con las aguas infiltracionales y dar lugar a una interfase caracterizada por poner en contacto aguas oxidantes con aguas reductoras. Algunos yacimientos de uranio pueden desarrollarse en esta interfase. Otras aguas pueden quedar atrapadas en su evolución, asociadas con gas natural o petróleo, o con uno y otro conjuntamente.

Los procesos petrogenéticos endógenos (plutonismo, vulcanismo y metamorfismo) pueden movilizar cantidades importantes de agua; agua que puede ser infiltracional o elisiónica y a la que la acción de los domos térmicos locales hace adquirir propiedades termales, pudiendo manifestarse en superficie como fuentes termales o sistemas hidrotermales submarinos. En estos procesos pueden existir *aguas juveniles*, es decir aguas de origen endógeno en el sentido riguroso del término, que nunca han estado vinculadas a la hidrosfera: ni son infiltracionales ni elisiónicas, sino que están relacionadas con la evolución del material interno del planeta.

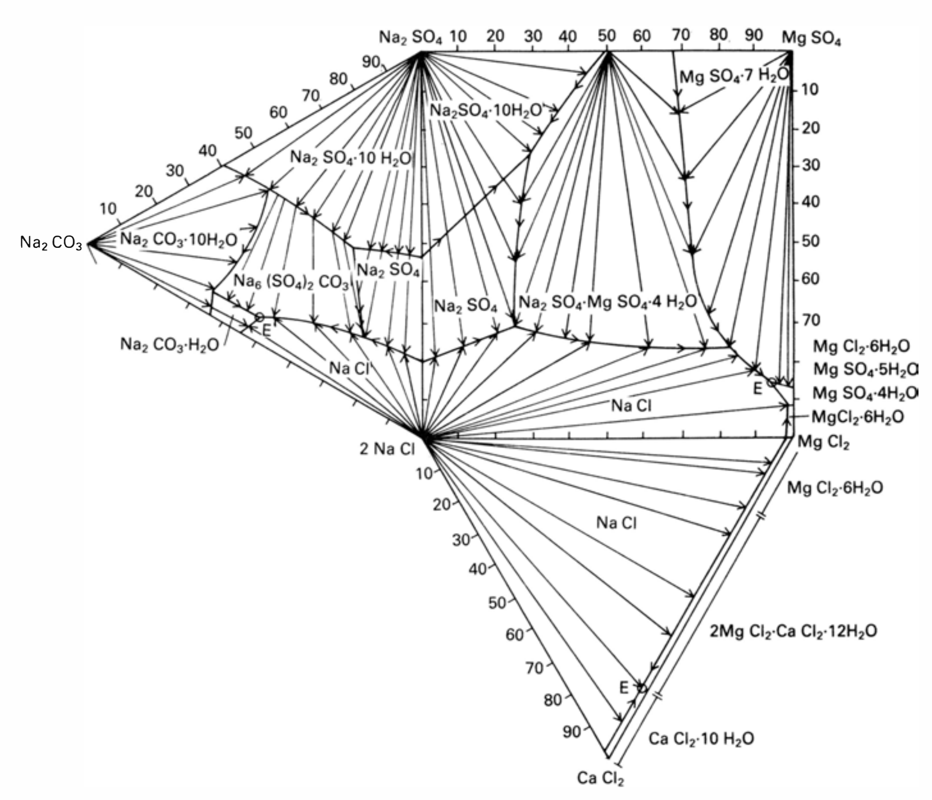
Aunque los procesos de génesis de una salmuera no son puros, se puede intentar una sistematización de los tipos de salmueras en función del proceso genético dominante. Las hay que han alcanzado su concentración en sales porque la evaporación natural supere las precipitaciones meteóricas y el aporte de las aguas dulces, de escorrentía o subterráneas. Este carácter está ligado fundamentalmente al clima y quizá más concretamente a la temperatura y al régimen de vientos. De una manera no tan clara están muy ligadas

algunas salmueras a determinados tipos de cuencas sedimentarias, aquellas que tienen una conexión restringida con aguas marinas y una escasez de recursos fluviales que las alimenten. Los salares andinos son un buen ejemplo de génesis de salmueras por causas climáticas y el golfo de Kara Bogaz, en el mar Caspio, lo es de génesis de salmueras por alimentación restringida, dada la ausencia de ríos importantes que desemboquen en dicho golfo, sumado todo ello a un clima árido con escasez de precipitaciones.

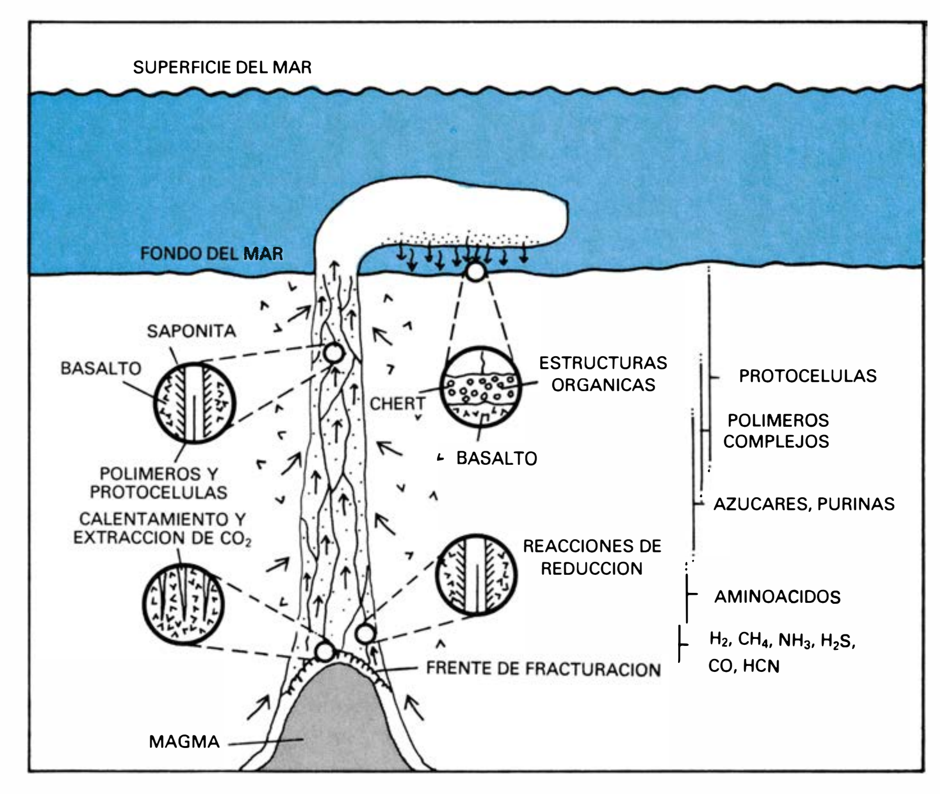
Para que la evaporación tenga lugar no es necesario que exista agua libre, sino que el agua que empapa el suelo pueda evaporarse; éste es el origen de la salinización de los suelos. Las salmueras formadas a partir de las aguas de imbibición de los suelos se denominan salmueras vadosas, recibiendo el nombre de lacustres las que se generan sobre las áreas continentales en lagos efímeros (playas) o permanentes, sin conexión aparente con aguas marinas. Las aguas marinas son quizá la salmuera mejor conocida y su evolución en cuencas marginales de pequeño espesor de agua (salinas), o en cuencas profundas con comunicación restringida con el mar abierto, han producido en el pasado gigantescos depósitos de sales naturales. La comunicación puede ser en la superficie o puede ser a través de las aguas subterráneas. Se comprende fácilmente que los datos isotópicos contribuyan a identificar las aguas marinas allí donde las conexiones no son evidentes.

Las salmueras infiltracionales son las generadas por la disolución de sales preexistentes: cuando las aguas infiltracionales locales alcanzan un nivel de sales, si se produce una surgencia, o simplemente si se extrae agua del acuífero mediante bombeo artificial, se genera una salmuera, tanto más concentrada cuanto más soluble es la sal y cuanto más tiempo permanece en contacto con la misma. Resulta más difícil imaginarse el largo recorrido de las aguas infiltracionales profundas, que alcanzan temperaturas bastante elevadas. En los últimos años la necesidad de explotar sales minerales y crear espacios útiles en el seno de las mismas, con fines estratégicos o de almacenamiento de sustancias radiactivas o peligrosas, ha permitido desarrollar la aplicación artificial de la infiltración como método de obtención de salmueras.

Las aguas del mar que se incorporan a los sedimentos forman salmueras que se caracterizan por un contenido en cloruros mayor que en sulfatos, más



10. TIPOS HIDROQUIMICOS DE SALMUERAS a partir de los diagramas de M. G. Valyashko. En el triángulo superior está el campo de las *salmueras sódicas*: son las aguas continentales o marinas muy continentalizadas; los cationes calcio y magnesio han sido eliminados como carbonatos y todavía existe exceso de carbonatos y bicarbonatos. El cuadrado representa las salmueras sulfatadas. En ellas existe un exceso de magnesio, que se traducirá en la formación de fases sulfatado-magnésicas; este proceso puede abortarse en sistemas naturales abiertos con entrada de aguas continentales carbonatadas formadoras de fases sulfatadas dobles o triples (Ca-Na-Mg-K). Las aguas marinas pertenecen a este grupo. Dada la composición de una salmuera, el diagrama permite prever la secuencia de sales por evaporación.



11. EL ORIGEN DE LA VIDA sobre la tierra parece estar relacionado con los intensos procesos de reducción del CO_2 de las aguas marinas en zonas de *rifting* o separación de placas; estas zonas presentan anomalias termales elevadas, junto con una fracturación fuerte que permite a las aguas marinas penetrar profundamente y adquirir temperaturas altas a la vez que incorporan cationes metálicos. En el fondo del mar Rojo se han encontrado salmueras calientes que dan lugar a depósitos de sulfuros metálicos. La hipótesis de Corliu, Baross y Hoffman (1981), en el sentido de que la "sopa primigenia" prebiológica está relacionada con salmueras calientes se recoge en este gráfico adaptado del original de dichos autores.

	Ca/Na	Mg/Ca	K/Na	Li/Na	HCO_3 /Cl	SO_4 /Cl	F /Cl	Br /Cl	I /Cl	B/Cl	SiO_2 ppm
AGUAS CONTINENTALES	*2,38*	0,27	*0,37*	*0,0006*	*7,49*	*1,44*	*0,03*	0,0007	0,0002	0,0015	13,1
AGUAS MARINAS	0,038	3,2	0,036	0,00001	0,0074	(0,14)	0,00007	0,0034	0,000003	0,000003	7,0
AGUAS ELISIONICAS: a) ClNa b) ClNa-Ca	0,04 (0,3)	0,4 0,15	0,015 0,02	0,0003 0,0002	0,02 0,001	0,000 0,0008	0,0002 0,00002	0,003 0,005	*0,002* 0,0008	*0,003* 0,0002	*30* 10
SALMUERAS CALIENTES (Mar Rojo)	0,051	0,17	0,023	0,00005	0,0004	0,0048	n.d.	*0,0079*	2×10^{-7}	n.d.	n.d.
SALTON SEA 300 °C	0,55	0,002	0,34	0,004	0,0006	0,000	0,000	0,000	n.d.	0,0004	400

12. LAS RELACIONES ENTRE IONES en las aguas naturales sirven para orientar la investigación de determinadas salmueras y contribuir a la interpretación de su génesis. Incluso pueden aplicarse a la interpretación de la génesis de fases salinas, pues éstas incorporan proporciones fijas de iones (factor de concentración) en relación al medio. Este cuadro refleja que las aguas continentales son las que presentan unas relaciones más ricas en calcio, potasio y litio frente a sodio, así como bicarbonatos, sulfatos y fluoruros frente a cloruros. Las aguas más ricas en cloruros frente a sulfatos son las aguas elisiónicas. La relación Br/Cl es más alta en las aguas marinas. El iodo y el boro presentan relaciones elevadas respecto al cloro en las aguas elisiónicas.

magnesio que calcio y mucho más sodio que potasio. Durante el enterramiento y por efecto de la materia orgánica, los sulfatos disminuyen, por reducción a sulfuros, quedando los cloruros prácticamente solos; el magnesio es adquirido por los carbonatos y por las arcillas con preferencia al calcio; el potasio se incorpora a la red de los filosilicatos con preferencia al sodio. En definitiva: el intercambio de iones en los silicatos, la disolución de los carbonatos y la maduración de la materia orgánica confieren a las aguas connatas una composición diferente de la inicial, con concentraciones mayores y con un contenido en iones pesados y algunos iones biófilos (ioduros y boratos) varias veces superior al del agua marina original.

El origen de las salmueras calientes es complejo, ya que a ellas se puede llegar por varias vías. Ahora bien, cualquiera que sea el tipo de agua calentada por una anomalía térmica local, la solubilidad de la mayoría de los minerales aumenta en condiciones de presión y temperaturas elevadas. La composición de la salmuera dependerá de la naturaleza del agua calentada y del tipo de rocas que atraviesa. El efecto más regular que se conoce es la relación molar potasio a sodio (K/Na), que depende de la temperatura.

Los aniones disueltos en las salmueras son el carbonato y bicarbonato, junto con el sulfato y el cloruro. Entre los cationes destacan el calcio, magnesio y sodio, y, en proporciones mucho más pequeñas, el potasio. Los aniones y cationes precipitados constituyen la base de la clasificación de las salmueras. Cuando los equivalentes de carbonato o bicarbonato presentes en la solución superan a los equivalentes de calcio y magnesio se habla de salmueras sódicas. (El calcio y el magnesio serán eliminados de la solución en una etapa precoz, ya que son fases insolubles, quedando el sodio como catión dominante, junto con otros cationes alcalinos minoritarios como son el potasio y el litio.) En cuanto a los aniones minoritarios, en las aguas sódicas pueden aparecer nitratos, sobre todo en ambientes muy oxidantes, por transformación de compuestos orgánicos nitrogenados, iodatos, boratos y otros. La sílice puede estar presente, tanto en solución iónica ($SiO_4H^-{}_3$) como coloidal ($SiO_4H^o{}_4$).

Las salmueras sódicas son alcalinas y por evaporación precipitan carbonatos sódicos (trona o gaylussita), sulfato sódico (thenardita o mirabilita) y cloruro sódico (halita). Genéticamente están relacionadas con la evolución, bajo climas áridos, de aguas continentales que drenan rocas plutónicas ácidas (granodioritas) y rocas volcánicas ácidas (serie calcoalcalina); en estas rocas predominan los silicatos alcalinos y, durante los procesos de meteorización, el sodio y el potasio fijan carbono atmosférico, que se incorpora así a las aguas en exceso.

Las salmueras en las que la cantidad de equivalentes de calcio y de magnesio es semejante a la de aniones carbonato y bicarbonato, de tal manera que se precipita todo el carbonato, pudiendo quedar todavía iones calcio o magnesio en exceso, o bien sólo ion magnesio, reciben el nombre de salmueras sulfatadas. Precipitan en etapas muy precoces, además de carbonatos de calcio, sulfatos de calcio (yeso), o bien carbonatos de calcio de alto contenido en magnesio sin que precipite yeso. Estas salmueras suelen ser neutras o ligeramente alcalinas. A este grupo pertenecen las aguas marinas estándar y las aguas infiltracionales que atraviesan series evaporíticas antiguas, marinas; constituye el grupo más abundante. Su mezcla con aguas continentales da una serie de términos intermedios entre salmueras marinas o paramarinas y salmueras sódicas continentales. De las salmueras sulfatadas se obtienen como precipitados normales sulfatos sódicos (mirabilita y thenardita), típicos de aguas sulfatadas en tránsito a aguas sódicas, y cloruro sódico, dominante en las salmueras de origen marino. Los precipitados de sulfato sódico-magnésico (bloedita) y sulfatos magnésicos (epsomita, hexahidrita) suelen hallarse en aguas sulfatadas infiltracionales y en zonas de oxidación de yacimientos de sulfuros.

Si una salmuera sulfatada se confina en un ambiente reductor, aspecto que está relacionado con la abundancia de materia orgánica y la falta de oxigenación, los aniones sulfato se reducen para dar aniones sulfuro. En ausencia de carbonatos también reducibles a metano y de aniones sulfato, los aniones cloruro dominan y nos encontramos ante aguas cloruradas. A partir de éstas se precipitan cloruros sódicos (halita), cloruros magnésicos (bischofita) y cloruros cálcicos (taquihidrita). Estas salmueras suelen tener *p*H neutro o ligeramente ácido. Pertenecen a este grupo las salmueras asociadas al petróleo y las salmueras del mar Muerto.

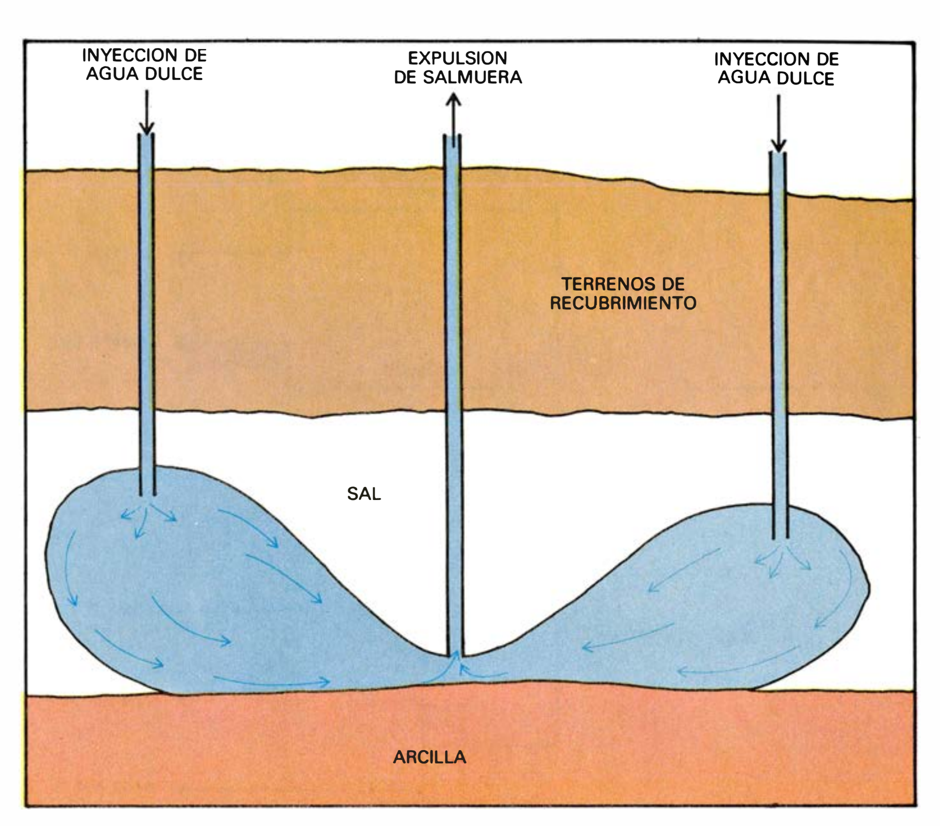
La relación Ca/Na es mucho mayor en las aguas continentales que en los demás tipos. La relación Mg/Ca es superior en las aguas marinas o relacionadas con ellas que en cualquier otra de las aguas conocidas. Las continentales y algunas termales presentan una relación K/Na elevada. El litio está muy vinculado a aguas termales, siendo un importante constituyente de las mismas. Por otra parte, las aguas continentales presentan, en relación a los cloruros, los más bajos valores, mientras que tienen los valores más altos de los aniones bicarbonato, sulfato y fluoruro. Las

aguas elisiónicas o diagenéticas son extraordinariamente ricas en los aniones borato, ioduro y bromuro.

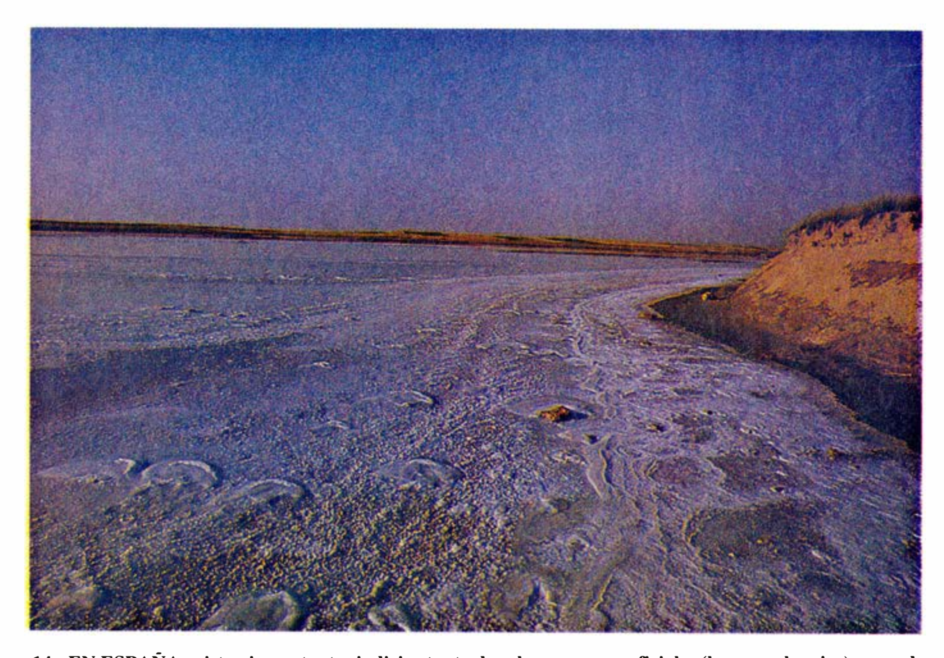
Los procesos naturales son abiertos. Cuando se establece la tipología hidroquímica de las salmueras, da la impresión de que se está hablando de sistemas cerrados. Esto lleva a una interpretación errónea de la génesis de sedimentos químicos a partir de salmueras. Hay mecanismos que permiten cambiar la evolución química de una salmuera; pueden modificarse las facies salinas previsibles de obtener por evaporación, a partir de la composición inicial. En los medios naturales a este fenómeno se le denomina metamorfismo de salmueras, y se ha estudiado para salmueras superficiales en contacto con aportes de aguas continentales (bicarbonatadas y cálcicas). Este mecanismo explica la génesis de sales dobles: sulfatos de sodio y calcio (glauberita), sulfatos de potasio y calcio (singenita) y sulfatos de potasio, magnesio y calcio (polihalita), que no podrían justificarse a partir de una secuencia normal de evaporación.

A la hora de rentabilizar el aprovechamiento industrial, el esfuerzo se aplica a los procedimientos siguientes: modificación de la presión parcial de gases como el oxígeno y el CO_2, que cambian el potencial redox y el *p*H de las aguas; la adición de compuestos que modifiquen el equilibrio de alguna de las fases en solución, en el sentido de aumentar su solubilidad o disminuirla; la adsorción diferencial de cationes o aniones sobre materiales con propiedades de superficie (ceolitas, minerales de arcilla, geles orgánicos e inorgánicos); variaciones de temperatura, favoreciendo, en su caso, la precipitación de sales criofílicas o termofílicas.

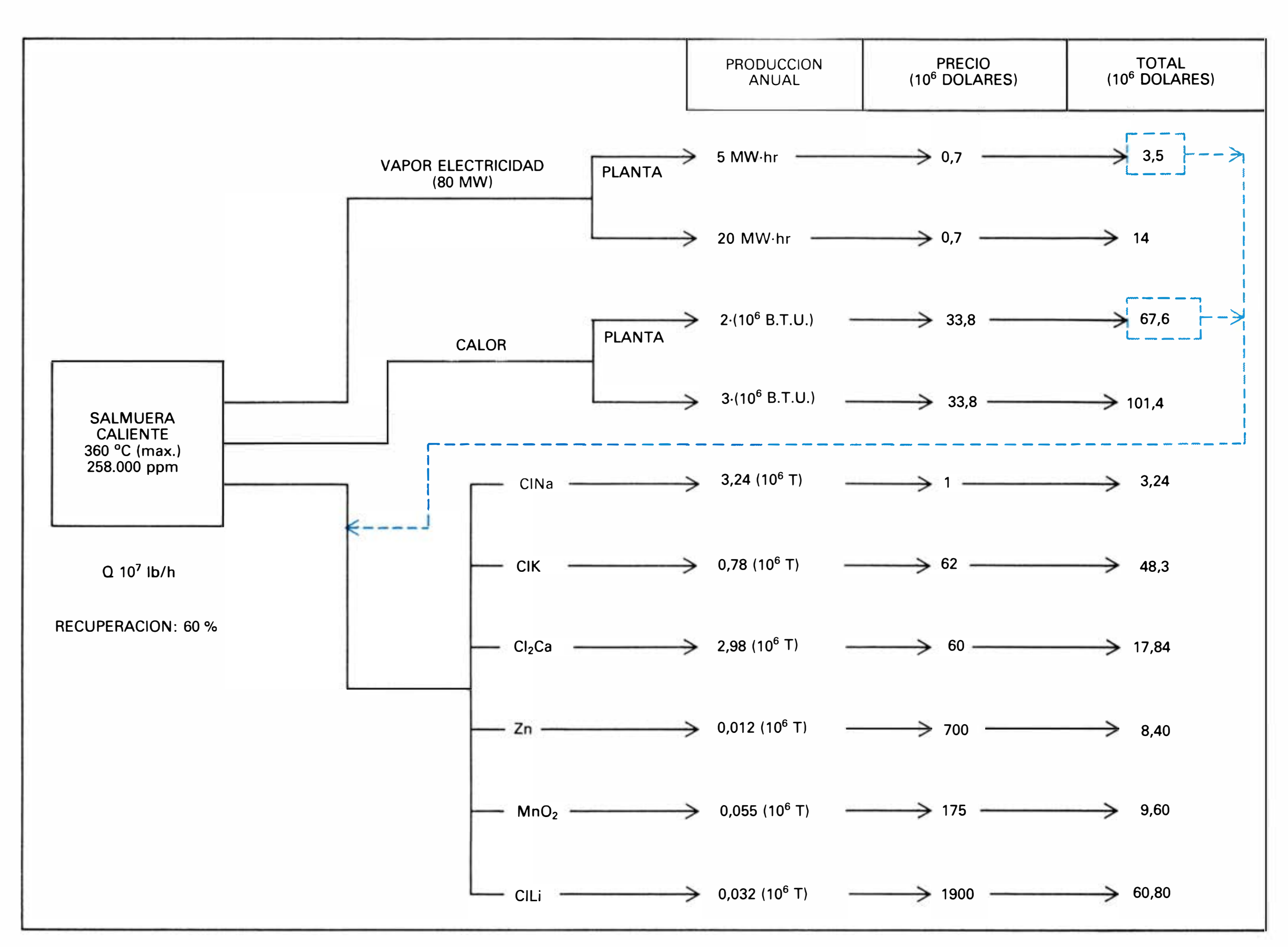
En la Europa del siglo XVI, la evaporación artificial, mediante calentamiento, suponía un avance tecnológico importante en la obtención de sales a partir de salmueras. La extracción en el lago Searles de carbonato sódico y sulfato sódico a partir de sus aguas es ya un proceso complejo, en el que interviene la precipitación por adición de sustancias y la redisolución de sales antes formadas. Para obtener carbonato sódico a partir de las salmueras del lago Searles se inyecta en las mismas CO_2 a presión, procedente de la calcinación de calizas, lo que provoca una sobresaturación, en bicarbonato sódico que precipita y que calcinado da lugar al producto final. Esta misma salmuera si se evapora, sin más, precipita en primer lugar halita (ClNa), para luego precipitar una sal doble: burkeita



13. CONSTRUCCION EN SALES de almacenes para sustancias peligrosas (residuos nucleares, hidrocarburos gaseosos, productos manufacturados y explosivos). Ha permitido poner a punto métodos de creación de cavidades mediante inyección de agua salada y extracción de la salmuera resultante. El proceso suele regularse mediante la utilización de una sustancia como gasoil que flota sobre la salmuera y condiciona que la disolución avance en horizontal, ya que en caso contrario podría desarrollarse una disolución en chimenea que haría peligrar la infraestructura de inyección y extracción. Este modelo se aplica también a la obtención de salmueras artificiales y permite explotar yacimientos salinos profundos.



14. EN ESPAÑA existen importantes indicios tanto de salmueras superficiales (lagunas, lavajos) como de salmueras subterráneas, de cantidad y tipología variable. No obstante, la panorámica actual de la explotación es artesanal y poco tecnificada, salvo en el caso de la explotación mixta de las salinas de Torrevieja y La Mata. En el valle del río Ebro existen lagunas de enorme extensión, muy ricas en sulfato sólido; en las llanuras manchegas abundan lagunas ricas en sulfatos sódicos-magnéticos y sulfatos magnésicos. Los lavajos de la cuenca del río Duero presentan contenidos apreciables en carbonato sódico. Sin duda, todas estas manifestaciones están relacionadas con zonas de descarga de flujos regionales subterráneos que atraviesan depósitos salinos. Pozos y manantitales de salmueras son asimismo comunes en relación con materiales salinos. En la fotografía se observa un aspecto de la laguna de Taray (Toledo) en julio de 1982. Las posibilidades de las aguas geotermales son menos conocidas que las de las superficiales.



15. PROYECTO DE APROVECHAMIENTO INTEGRAL de la salmuera del Salton Sea, con un caudal de explotación de 5000 metros cúbicos por hora y con una recuperación del 60 por ciento. Se indican las producciones anuales y los valores de la producción vendible, así como la parte de la energía total consumida en los procesos de obtención de las sales. [Datos tomados de E. F. Wahe: *Geothermal energy utilization*; John Wiley and Sons. N. Y. 1977].

($CO_3Na_2 \cdot 2SO_4Na_2$); este mineral se redisuelve y, al enfriar la solución, precipita mirabilita ($SO_4Na_2 \cdot 10H_2O$), que por calentamiento dará SO_4Na_2 anhidro, que constituye un producto comercial. La solución enfriada puede a su vez tratarse para obtener fases minerales carbonatadas. Vemos por tanto cómo una salmuera constituye un yacimiento de baja ley que una tecnología adecuada puede convertir en un recurso económico, no sólo de las sales de los metales alcalinos y alcalinotérreos, sino incluso de otros metales (elementos y compuestos). Llaman la atención las diferentes patentes desarrolladas basadas en la adsorción diferencial en geles de algunos cationes metálicos, que aplicadas a salmueras ricas en estos metales pueden dar lugar a concentraciones sorprendentes.

El aprovechamiento de las salmueras naturales ha tenido un competidor en los productos sintéticos; pero la crisis energética ha vuelto a plantear las posibilidades de aquéllas. Al igual que en todos los recursos naturales, el aspecto económico de los mismos se debe plantear en las etapas de prospección, valoración y en su caso explotación. Las salmueras tienen muy diferentes orígenes; las superficiales, de acceso directo, son fáciles de prospectar. Las subterráneas constituyen un subproducto de la utilización de otros recursos, como las salmueras connatas de los hidrocarburos y las geotermales de los campos de energía geotérmica. En todos los casos han de plantearse dos cuestiones básicas: naturaleza y contenido en sustancias en solución, así como caudal disponible de salmueras para unas determinadas obras de infraestructura extractiva (pozos, canales, superficie de evaporadores, etcétera). Otro factor básico a considerar será el grado de homogeneidad en el espacio y en el tiempo del producto, es decir, que una vez diseñado un modelo de explotación, éste no podrá ser sustancialmente variado; por ello es necesario que la naturaleza y concentración de las soluciones no varíe excesivamente. El aprovechamiento debe plantearse utilizando al máximo sus propiedades específicas: su alto índice de refracción, que las hace funcionar como buenos almacenes de radiación solar, su bajo calor específico, su alta densidad, su conductividad eléctrica y otras. En lugares donde la evaporación potencial sea mayor que las precipitaciones, climas áridos, la evaporación natural constituye un buen método de enriquecer la concentración inicial, y sólo requiere una inversión inicial en la construcción de los evaporadores.

Las salmueras son muy sensibles a los cambios térmicos durante su transporte, sobre todo cuando se trabaja con soluciones próximas a saturación, pudiendo provocar encostramientos en los salmueroductos. Ejercen también efectos corrosivos sobre los metales. En definitiva se puede decir que las salmueras naturales son un recurso, en la gran mayoría de los casos, subeconómico, al que la tecnología y la imaginación pueden llegar a hacer económico. Las salmueras han sido importantes en la historia de la humanidad y la tecnología del año 2000 volverá a plantear su uso intensivo. Y este futurible se hará realidad debido a su carácter de recurso casi renovable y al bajo impacto medioambiental que representa su explotación, así como su posible uso combinado con la explotación de otros recursos y la baja peligrosidad para el ser humano que representa su explotación.

Centauro-A: la galaxia activa más cercana

Las galaxias activas irradian hasta un millón de veces más energía que las normales. A través de las observaciones de Centauro-A se estudia cómo se emite dicha radiación

Jack O. Burns y R. Marcus Price

Entre los objetos de mayores proporciones, rodeados de misterios, que se observan en el universo están las galaxias activas. Constituyen sólo un pequeño porcentaje de cuantas conocemos; pueden emitir, cada una, hasta un millón de veces más energía, en forma de radiación electromagnética, que cualquier galaxia ordinaria. Las emisiones de una galaxia activa se extienden a muy diferentes frecuencias: rayos gamma, rayos X, radiación visible, radiación infraroja y ondas de radio. Gran parte del trabajo reciente de observación se ha realizado en frecuencias de radio, donde la emisión de una galaxia activa puede multiplicar por 100.000 la intensidad de otra ordinaria.

La galaxia activa más cercana a la nuestra es Centauro-A, que se encuentra a unos 15 millones de años-luz. El perfil visible de Centauro-A tiene un tamaño de unas decenas de miles de años-luz. Extendiéndose y formando ángulo con el plano central de la imagen visible, se distingue una radioenvoltura alargada, de unos tres millones de años-luz de un extremo a otro. Si la región radioemisora de Centauro-A fuera visible para un observador instalado en la superficie terrestre, le parecería un objeto unas 20 veces mayor que la Luna.

¿Cómo puede crearse y mantenerse la enorme radioestructura de una galaxia activa del tipo de Centauro-A? La combinación de observaciones de rayos X, de ondas de radio y de luz visible empieza a suministrar a los astrónomos copiosa información sobre la organización de Centauro-A. Una "motobomba" expulsa desde el centro de la galaxia el material de la radioestructura, como una manguera con su chorro de agua. De esta motobomba sale un haz delgado de radioplasma: gas altamente ionizado, cuyos principales componentes son electrones (que se mueven con velocidad cercana a la de la luz), campos magnéticos y, posiblemente, protones. Mientras los electrones viajan desde el centro de la galaxia hacia fuera emiten radiación en un amplio rango de frecuencias. Así, pierden energía y si ésta no les fuera suministrada de manera continua pronto pararía su movimiento. En su camino desde la motobomba hacia el exterior, el haz atraviesa tres radiolóbulos ordenados según tamaño creciente. El lóbulo medio coincide con una serie de filamentos ópticos, regiones delgadas que emiten radiación en el rango visible de frecuencias. Se acaba de demostrar que los filamentos son regiones de nacimiento de nuevas estrellas.

Los últimos diez años han registrado un notable progreso en nuestro conocimiento de las galaxias activas. Pero seguimos sin entender algunas características fundamentales. No se conoce con certeza la naturaleza de la motobomba que conduce el haz de plasma, y únicamente pueden hacerse hipótesis sobre los pormenores de su funcionamiento. Además, hay varias posibles explicaciones acerca de cómo son acelerados los electrones del haz mientras avanzan desde el núcleo galáctico hacia el exterior. Ateniéndonos a los datos, no nos está permitido decidir cuál es la explicación correcta. Tampoco sabemos cómo interaccionan las componentes radio y óptica de la galaxia en las regiones donde se forman estrellas jóvenes. Observaciones recientes de radioondas y rayos X de Centauro-A han facilitado grandemente nuestra segura incursión en los procesos que se desarrollan en estas galaxias activas. Por un lado, la nueva información nos revela una Centauro-A más familiar; por otro, sin embargo, las observaciones han descubierto procesos físicos tan poderosos y extensos que convierten esas galaxias en fenómeno aún más extraño y, a la vez, más maravilloso.

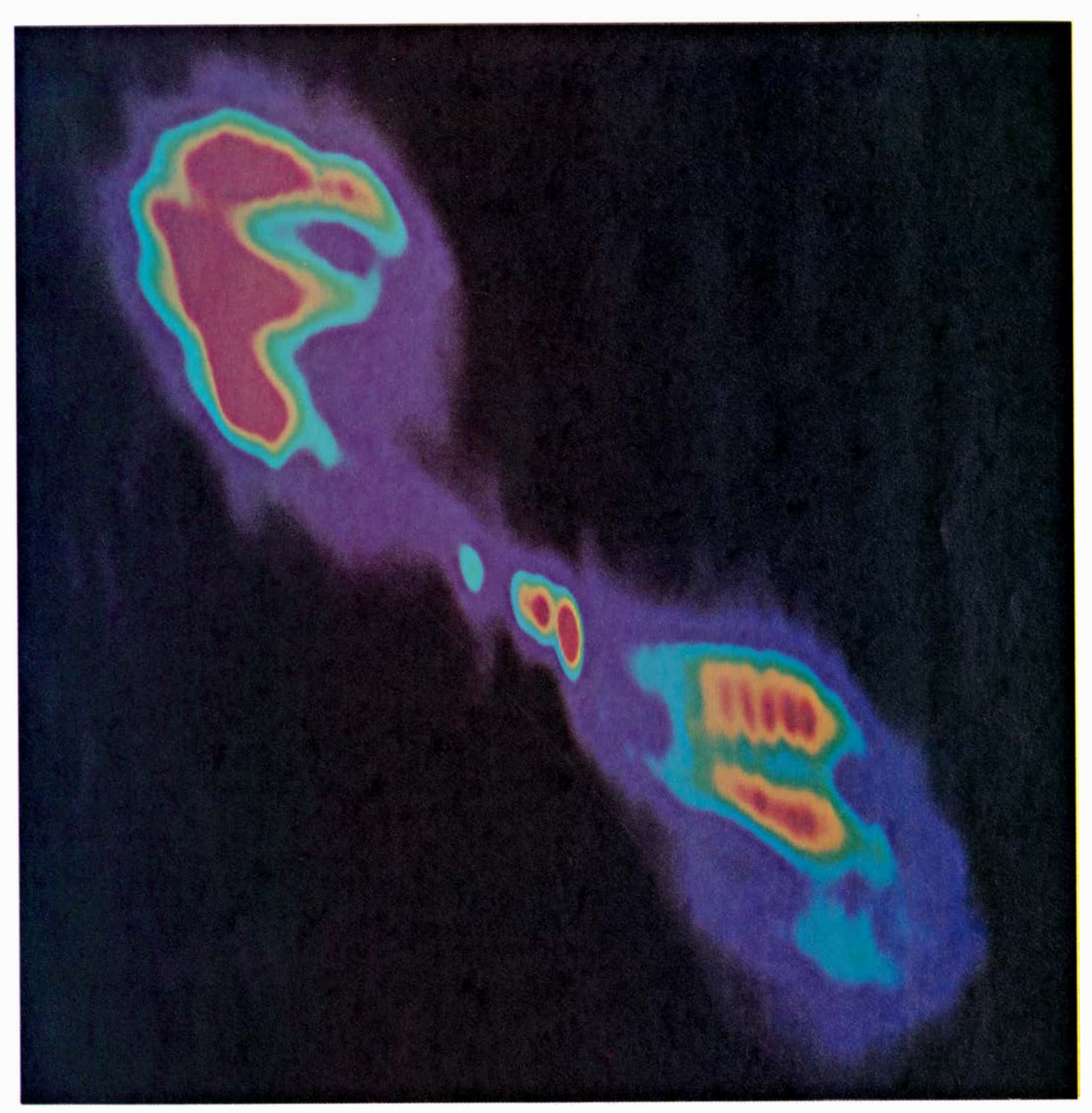
A Centauro-A, catalogada como NGC 5128 en las observaciones en el espectro visible, se le dio este nombre por ser la fuente de radio más luminosa en la constelación del Centauro. Vista desde el hemisferio Norte, esta constelación ocupa una posición meridional. Vista en el hemisferio Sur, desde Australia, está sobre nuestra cabeza. Desde Nuevo México, donde se han realizado muchas de las recientes observaciones en radiofrecuencia (utilizando el radiotelescopio Very Large Array [VLA] perteneciente al Observatorio Nacional de Radioastronomía de los Estados Unidos), la galaxia está solamente unas pocas horas al día sobre el horizonte; su elevación máxima es de unos 14 grados.

Centauro-A es una de las fuentes de radioondas más potentes de todo el firmamento. La amplitud de emisión y su proximidad con respecto a nuestra galaxia permiten observar procesos y ver estructuras que pueden estar presentes en todas las galaxias activas, si bien no pueden detectarse en las más distantes. En los últimos años, y como resultado de todo ello, los astrónomos han prestado gran atención a Centauro-A. Se ha estudiado con la mayoría de los más sensibles detectores de radiación astronómica, incluidos, además del VLA, el telescopio óptico de 4 metros enclavado en Chile en el Observatorio Interamericano de Cerro Tololo, el telescopio Angloaustraliano ubicado en Siding Spring (Australia) y el Observatorio Einstein, de rayos X.

En los trabajos sobre Centauro-A, gran parte de la radiación detectada se origina en tres procesos físicos distin-

tos. En todos ellos se emite radiación cuando un electrón pierde energía. Pero en cada proceso varían los pormenores del cambio de energía de los electrones y, también, las características de la radiación emitida. En el primer proceso, la radiación incidente suministra energía a las nubes de gas interestelar. Los átomos de gas de tales nubes absorben energía radiante y, en consecuencia, se ionizan (es decir, ganan o pierden electrones). Cuando un electrón de la nube pasa cerca de un ion sufre una desviación y una aceleración; la mayoría de los iones son protones, núcleos de hidrógeno privados de su único electrón. Cuando la trayectoria del electrón cambia, éste emite un fotón, cuanto de energía electromagnética. A esta radiación se la denomina radiación térmica, ya que su espectro depende estrechamente de la temperatura del gas.

Se puede describir la radiación electromagnética como flujo de fotones o como tren de ondas. En la primera descripción, se caracteriza a la radia-

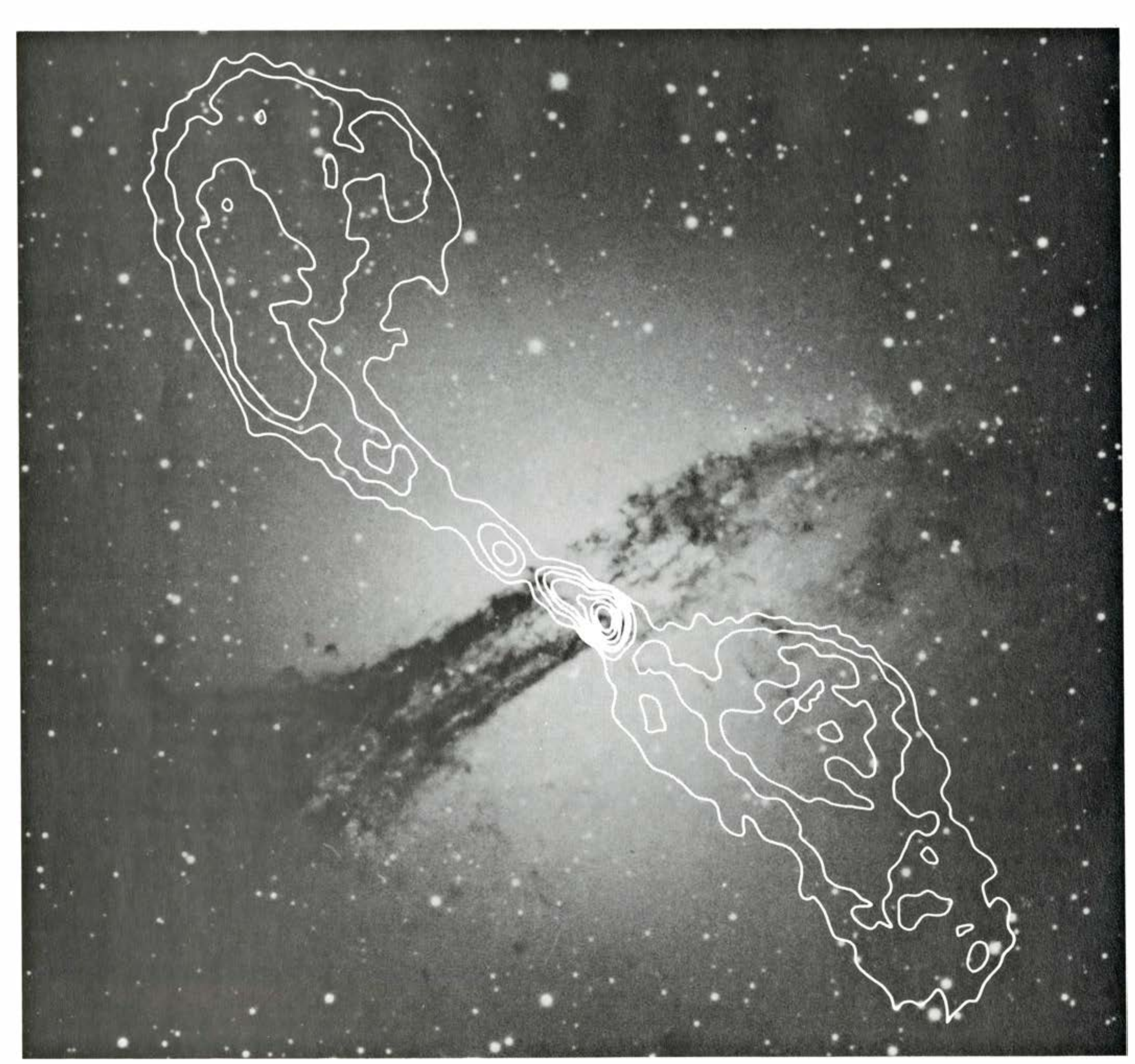


1. RADIOLOBULOS INTERNOS de Centauro-A, representados en un mapa de falsos colores de las radioemisiones de la región central de la galaxia, catalogada en el visible como NGC 5128. El código de colores en este mapa es tal que el rojo representa la radioemisión más intensa, mientras que en azul tenemos la menos intensa. En ésta y otras imágenes de la galaxia en las próximas páginas, el Norte cae arriba y el Este a la izquierda. Los lóbulos internos son las estructuras simétricas redondeadas alineadas en un eje que va del nordeste al sudoeste, con una longitud de 60.000 años-luz. Además forman la radioestructura más pequeña y más interna de una serie de ellas que se extienden desde el centro de la galaxia hacia cada lado. Al canal estrecho entre el centro de la galaxia y el lóbulo interno norte se le llama chorro. Se le supone un haz de plasma, es decir, de gas altamente ionizado. El plasma es expulsado desde el núcleo galáctico, representado por el mayor de los óvalos rojos entre los lóbulos internos. Como en otras galaxias activas, el plasma emite en un rango muy ancho de frecuencias. Las radioemisiones fueron observadas por uno de los autores (Burns), Eric D. Feigelson y Ethan J. Schreier con el radiotelescopio VLA de Nuevo México. El mapa se hizo con ayuda de un sistema de proceso de imágenes por ordenador, desarrollado por Dennis Ghiglia.

ción con la energía de los fotones; en la descripción ondulatoria, la correspondiente propiedad es la frecuencia, que guarda una relación directa con la longitud de onda. Cuanta más energía tenga un fotón, más alta será la frecuencia de las ondas y más corta su longitud de onda. En la radiación térmica la energía de los fotones emitidos, o la frecuencia de las ondas emitidas, se distribuye de forma continua, esto es, dibujan una curva suave en su representación gráfica. Sin embargo, la radiación tiende a ser más intensa en un rango de frecuencias bastante pequeño. Las ondas están polarizadas aleatoriamente, es decir, el plano de las oscilaciones electromagnéticas carece de orientación definida.

En el segundo proceso, el electrón es capturado y acelerado siguiendo una órbita helicoidal alrededor de una línea de fuerza del campo magnético. A medida que el electrón se acelera, emite radiación electromagnética en dirección tangencial a su trayectoria. A este tipo de emisión se la denomina radiación de sincrotrón; igual que la radiación térmica, presenta un espectro continuo de frecuencias, extendiéndose sobre un intervalo más ancho que aquélla. Además, en la radiación de sincrotrón, las ondas están polarizadas: las componentes del vector campo eléctrico de todas las ondas están alineadas en el plano perpendicular a la línea de fuerza magnética alrededor de la cual gira el electrón.

En el tercer proceso, se emite un fotón cuando un electrón salta, dentro de un átomo, de un nivel permitido de energía a otro más bajo. Cuando un átomo de la nube de gas absorbe energía de su entorno, un electrón de dicho átomo puede saltar a un nivel de mayor energía. Una vez que ha subido, el electrón tiende a regresar rápidamente al nivel de energía original. En la transición de bajada se libera un fotón de igual energía que la del fotón absorbido. Los fotones emitidos así muestran



2. IMAGEN OPTICA y radiocontorno de Centauro-A, dispuestos perpendicularmente. Esta figura combina los radiocontornos de los lóbulos internos con una fotografía hecha en el visible. El área visible de NGC 5128 tiene una forma atípica con detalles de galaxias espirales y elípticas. El gran círculo luminoso se parece mucho al cuerpo de una galaxia elíptica. La banda oscura que corta al círculo es una franja de polvo, gas y asociaciones de estrellas con forma de disco, que nos recuerda las galaxias espirales. El disco se dobla en sus bordes en direcciones opuestas. A pesar de que la elipse gira muy lentamente, si lo hace, la franja de polvo gira de manera estacionaria. La radioestructura emerge desde el centro de la franja a lo largo de su eje de rotación.

un espectro discreto de frecuencias, en que cada una de ellas corresponde a una transición particular dentro del átomo. Del continuo que representa el espectro electromagnético total, la radiación emitida aparece como un grupo de líneas brillantes conocidas por líneas de emisión atómica. La radiación de estas líneas de emisión no está polarizada.

Como veremos más adelante, la radiación de sincrotrón ha resultado ser de gran importancia para la comprensión de la física de Centauro-A. Sin embargo, las primeras observaciones de la galaxia se hicieron en el rango visible, lo que no debe sorprendernos, ya que NGC 5128 es el séptimo objeto más luminoso fuera de nuestra galaxia. De las abundantes observaciones ópticas, que comienzan a principios del siglo XIX, ha llegado a establecerse que la parte visible de Centauro-A tiene una estructura muy atípica, con características de galaxias espirales y elípticas a la vez.

La parte principal de NGC 5128 es una elipse brillante que subtiende un ángulo de siete minutos de arco. Como la distancia estimada es de 15 millones de años-luz, cada minuto de arco representa unos 5000 años-luz y, por tanto, la elipse tiene una longitud de unos 35.000 años-luz. Su masa se ha estimado en unos 300.000 millones de veces la solar. Tal masa y dicho tamaño son propios de los sistemas elípticos.

Lo que ya no es usual es la presencia en medio de la elipse de una ancha franja de absorción, una franja de polvo. En las imágenes ópticas la franja de polvo aparece a modo de banda oscura alineada con el eje menor de la elipse. La franja tiene la forma de un disco con sus bordes deformados en direcciones opuestas.

La franja de polvo consta, además del mismo, de asociaciones de estrellas y regiones ocupadas por hidrógeno atómico ionizado. Se designa H I al hidrógeno neutro, no ionizado y H II, al ionizado. Por tanto, las regiones de hidrógeno ionizado que se observan en la franja de polvo de NGC 5128 serán regiones H II. En ellas, el gas de hidrógeno tiene una temperatura de decenas de miles de grados Kelvin (grados Celsius por encima del cero absoluto). El gas se calienta por la radiación proveniente de las jóvenes y calientes estrellas azules que se han formado recientemente en el interior de la nube de gas. Las regiones H II muestran una serie de líneas de emisión características.

El cuerpo principal de Centauro-A apenas gira, como en la mayoría de sistemas elípticos. Sin embargo, la franja de polvo está girando uniformemente a la velocidad propia de las galaxias espirales o tipo disco. La franja de polvo y la elipse están centradas en el mismo punto; el potencial gravitatorio de esa enorme elipse domina al sistema compuesto por componentes. Tal configuración híbrida, extraordinariamente rara, podría muy bien guardar relación con la violenta actividad que subyace a la radioestructura.

Antes de volver a la observación de ondas de radio y de rayos X, que son el centro del actual interés en Centauro-A, conviene hacer notar otro rasgo óptico inusual. Placas fotográficas de larga exposición hechas por J. A. Graham, en Cerro Tololo, muestran la existencia de una notable región luminosa alargada en el cuadrante nordeste de la galaxia. Las placas se impresionaron con luz procedente de la línea de emisión atómica que corresponde a la recombinación de un electrón y un protón en las regiones H II, conocida por transición H-alfa. Las placas muestran que a una distancia de 50.000 años-luz del centro de la galaxia hay un filamento lineal que incluye tres inmensas regiones H II alineadas radialmente desde el centro hacia fuera de la galaxia. Más lejos, a una distancia de 130.000 años-luz, hay regiones que producen líneas de emisión e incluyen tanto filamentos como gas difuso. Por estudios espectroscópicos sabemos que la excitación del gas decrece con la distancia al núcleo de la galaxia; ello sugiere que la fuente que excita al gas está en el centro de la misma.

Cerca de algunas de estas alargadas regiones H II parece haber series encadenadas de brillantes estrellas jóvenes y azules; estas series fueron observadas por vez primera por Victor M. Blanco y sus colegas en Cerro Tololo. Si son realmente estrellas de una edad menor a diez millones de años, hubieron de formarse cerca de donde son hoy observadas, en vez de hacerlo en el centro de la galaxia para ser luego expulsadas. Las estrellas están a unos 65.000 años-luz del núcleo galáctico y, aún a una alta y poco frecuente velocidad de 1000 kilómetros por segundo (muchas veces mayor que la velocidad necesaria para escapar del campo gravitatorio producido por el cuerpo principal de la galaxia), las estrellas no habrían tenido tiempo de viajar desde el centro de la galaxia hasta donde están ahora.

La primera radioobservación detallada de Centauro-A comenzó en 1961, cuando entró en servicio el radiotelescopio de 64 metros del Observatorio Nacional de Radioastronomía de Australia. Uno de nosotros (Price) tuvo la fortuna de trabajar, en el último ciclo de su formación, con J. G. Bolton y B. F. C. Cooper, ambos de dicho observatorio, utilizando el nuevo radiotelescopio, que era entonces el primer instrumento de radioastronomía, para estudiar Centauro-A. No tardó en descubrirse que la radioemisión de la galaxia estaba polarizada, lo que abogaba por un proceso de sincrotrón.

Ulteriores trabajos mostraron que el plano de polarización de la emisión gira cuando las ondas de radio atraviesan las nubes de electrones de rayos cósmicos y los campos magnéticos débiles existentes en el espacio interestelar de nuestra galaxia. Tal rotación, llamada efecto Faraday, es proporcional al cuadrado de la frecuencia de la radiación. Aunque el efecto Faraday había sido predicho teóricamente, las observaciones de Centauro-A realizadas con el radiotelescopio australiano marcan un hito: es la primera vez que el efecto es detectado experimentalmente sin ambigüedades.

Así pues, las primeras radioinvestigaciones de Centauro-A tuvieron dos resultados notables. Primero, se confirmó la existencia de un campo magnético interestelar en nuestra propia galaxia. Segundo, la confirmación de la hipótesis de que las emisiones de radio de NGC 5128 eran consecuencia de un proceso de sincrotrón, un dato capital en la comprensión de las galaxias activas.

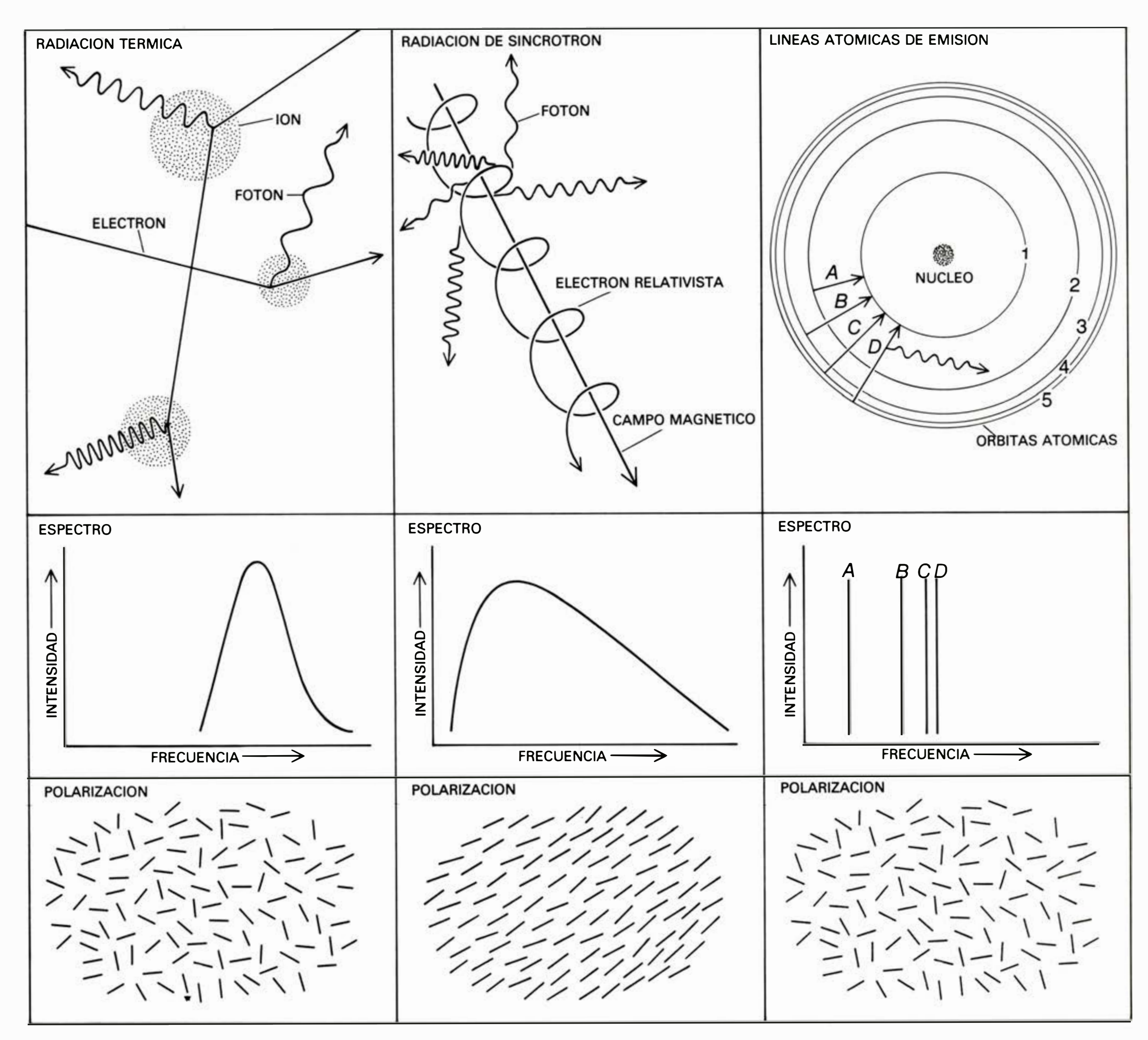
Las radioobservaciones australianas proporcionaron, además, el primer mapa fidedigno de la radioestructura de Centauro-A. Se encontró que la extensión de la radiorregión, de 2,7 millones de años-luz, cubre unos diez grados en el cielo. Cuando al comienzo de los años sesenta se radiocartografió por vez primera Centauro-A, constituía el mayor objeto astronómico hasta entonces observado. En los últimos diez años, se han encontrado fuentes de radio superiores, pero Centauro-A persiste en el elenco de los mayores objetos astronómicos conocidos.

Las extensas regiones emisoras de radiofrecuencia de cada lado del centro galáctico cambian de dirección con su distancia al núcleo. En la vecindad de éste, el alineamiento es nordeste-sudoeste; las regiones externas están alineadas casi norte-sur. El desplazamiento en dirección, de unos 65 grados,

desde el centro de la galaxia hasta el extremo de la radioenvoltura débese a una curvatura continua a lo largo de la longitud de la radiorregión. En virtud de lo cual, la radioestructura global adquiere la forma de una inmensa *S*. La curvatura indica que el flujo de material hacia la radioenvoltura se perturba continuamente. La perturbación podría obedecer a la precesión (balanceo como el de una peonza) de la motobomba nuclear o al efecto del gas galáctico exterior al núcleo que presiona contra el haz de plasma.

Entre el núcleo galáctico y el extremo de la radioenvoltura existen una gama de complejas radioestructuras. El tamaño de estas subestructuras aumenta con la distancia al núcleo, siendo las mayores un millón de veces más grandes que las pequeñas. Sólo recientemente, trabajando con el VLA, Eric D. Feigelson, de la Universidad de Pennsylvania, Ethan J. Schreier, del Space Telescope Science Institute, George W. Clark, del MIT, y nosotros hemos podido resolver la naturaleza de las radiocomponentes.

En un interferómetro de ondas de radio, el VLA es un ejemplo, las señales de radio procedentes de una sola fuente se reciben simultáneamente con un conjunto de varias antenas. En el VLA hay 27 antenas colocadas a lo largo de tres brazos que convergen. Comparando las relaciones de fase de las ondas que llegan a los receptores, puede obtenerse una representación bastante completa de la radiofuente. Con el VLA hemos logrado una resolución de un segundo de arco. Teniendo en cuenta la distancia a la que se en-

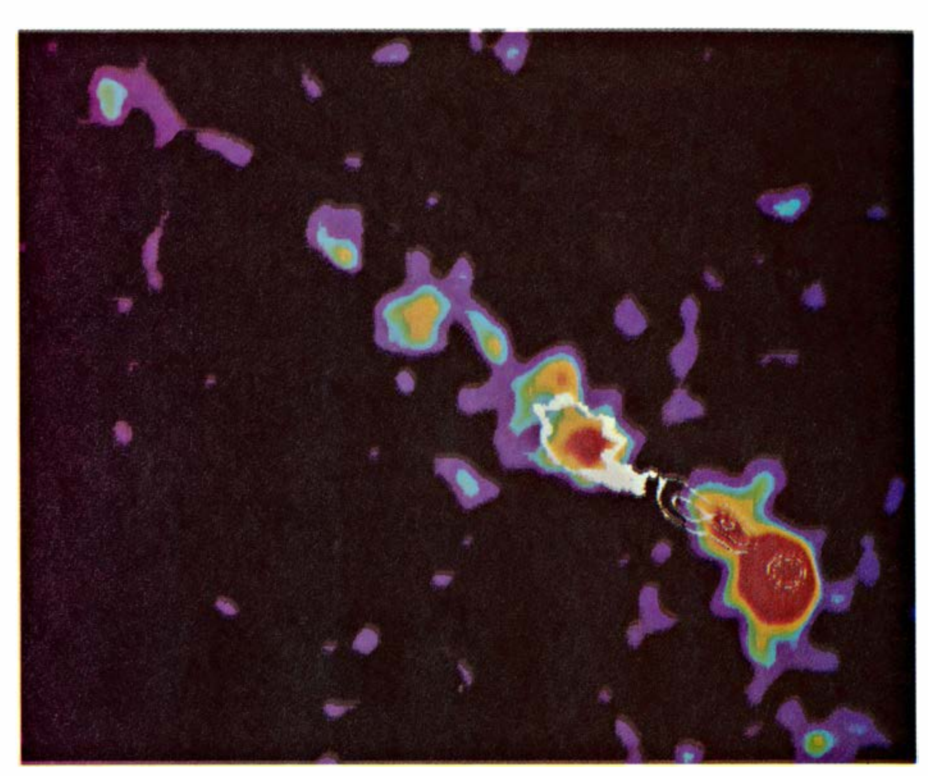


3. LA RADIACION EMITIDA POR CENTAURO-A puede originarse en tres procesos físicos distintos. En todos ellos un electrón pierde energía y el mecanismo a través del cual ésta se disipa determina la frecuencia e intensidad de la radiación. En la radiación térmica (*izquierda*) los electrones cambian de trayectoria y se aceleran al pasar cerca de iones (núcleos de hidrógeno, o protones) de una nube de gas caliente. Cuando un electrón se acelera emite un fotón, o cuanto de radiación, dotado de una energía, una frecuencia y una longitud de onda. Las frecuencias de la radiación térmica se distribuyen de manera continua, pero la emisión es más intensa en un rango bastante estrecho; la radiación térmica no está polarizada. En la radiación de sincrotrón (*centro*) los fotones son emitidos por electrones relativistas, electrones con una velocidad muy cercana a la de la luz. Los electrones relativistas se aceleran en trayectorias helicoidales alrededor de líneas de fuerza de un campo magnético; los fotones son emitidos en dirección tangente a la trayectoria. Esta radiación, que se emite en un gran rango de frecuencias, está polarizada. Otra manera de emitir el electrón un fotón (*derecha*): pasando de un nivel de energía permitido a otro de más baja energía. El espectro de esta radiación atómica consiste en una serie discreta de líneas y no está polarizada.

cuentra Centauro-A, dicha resolución equivale a unos 80 años-luz, comparable a la que se obtiene para la radiación visible en las placas fotográficas obtenidas con grandes telescopios.

Yendo desde los bordes de la radioestructura hacia el núcleo se aprecia una clara asimetría. Entre 70.000 y 130.000 años-luz al nordeste del núcleo galáctico, hay un gran radiolóbulo intermedio aproximadamente circular. No se ve ningún lóbulo medio comparable en la posición correspondiente al sudoeste. El lóbulo intermedio es una región de intensa emisión que se distingue de la difusa radioenvoltura exterior. La posición del lóbulo medio coincide con la de los filamentos visibles donde se contemplan las estrellas jóvenes. Como explicaremos después, se solapa también con una zona recién descubierta que emite rayos X.

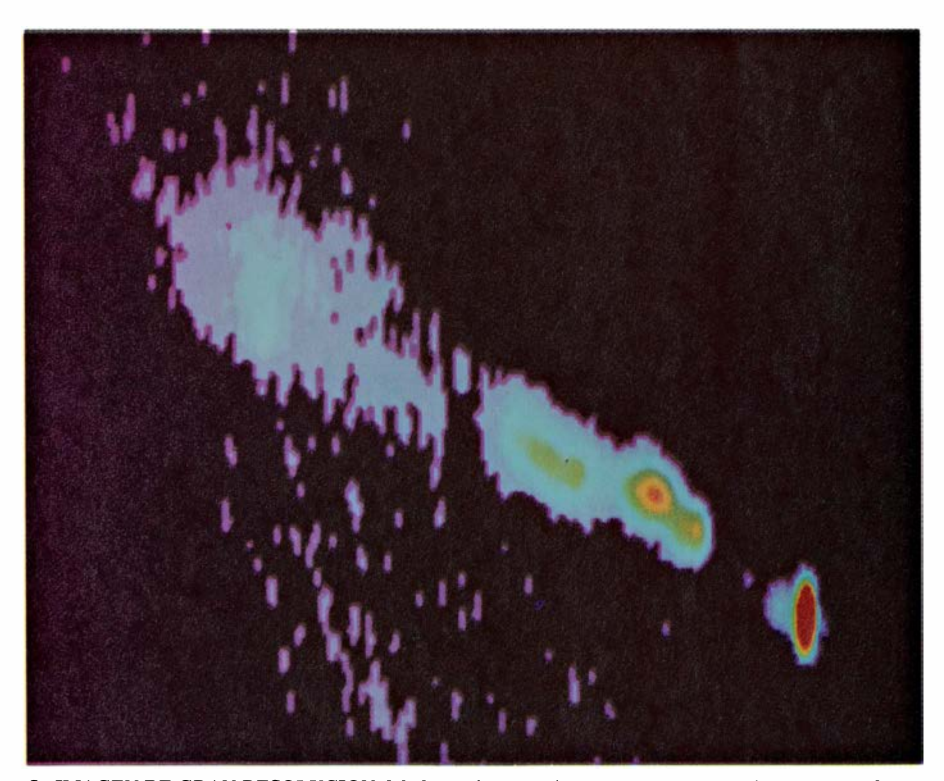
Cercana al núcleo existe una pareja de radiolóbulos internos, uno a cada lado. Los lóbulos internos, que son diferentes de los que se distinguen en la envoltura externa, se hallan ubicados simétricamente a 30.000 años-luz del centro de la galaxia. Con un diámetro de 3500 años-luz, son bastante más pequeños que el lóbulo medio. Aún no se ha observado radiación visible o rayos X asociados con los lóbulos internos.



4. CHORRO INTERNO, haz de plasma altamente colimado que lleva electrones energéticos procedentes del núcleo de Centauro-A hacia el lóbulo interno Norte. El núcleo galáctico es la región compacta roja en la parte inferior derecha, con contornos blancos superpuestos; el chorro se extiende hacia el nordeste desde el núcleo y tiene una longitud de 3000 años-luz. La imagen producida por el ordenador combina datos de radioondas y de rayos X recogidos por Feigelson y sus colegas usando el satélite-observatorio Einstein. Las áreas coloreadas corresponden a emisón en rayos X y los contornos blancos a radioemisión. El conjunto de las emisiones de radioondas y de rayos X sugieren que la radiación emitida por el plasma se debe a un proceso de sincrotrón. No se observa ningún chorro similar en el lado suroeste del núcleo.

Conectando lóbulo interno del norte y núcleo se aprecia un canal de emisión de radioondas y rayos X, estrechamente colimado, llamado chorro. Tales chorros, observados en muchas radiogalaxias, podrían constituir muy bien canales que transportaran electrones recién activados y campos magnéticos desde la motobomba central hasta el lóbulo interno. Si ello es verdad, el chorro sería un canal poco eficaz y con fugas, puesto que parte de su energía se irradia hacia fuera en forma de emisión de sincrotrón.

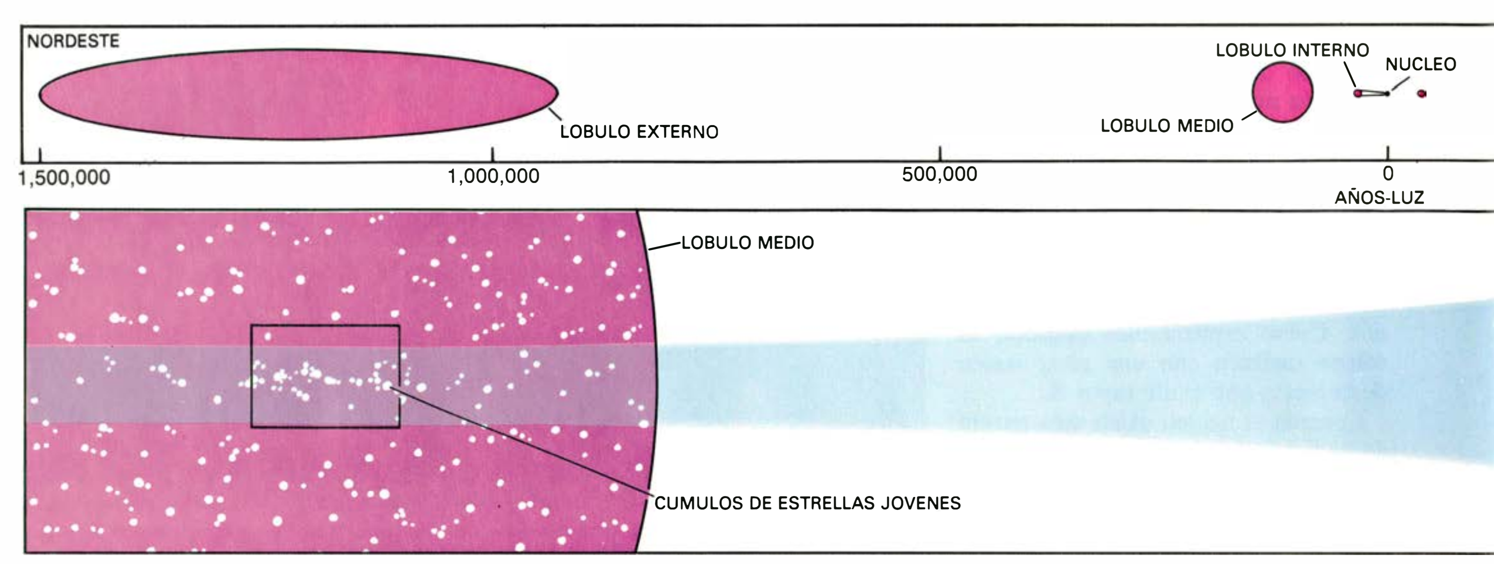
Quizá la asimetría más sorprendente de Centauro-A sea que únicamente la parte norte de la radioenvoltura tiene un chorro; no se observa ninguna estructura similar en la parte sur. Muchas radiogalaxias tienen un chorro que sale del núcleo galáctico y, en algunas de ellas, hacia un solo lado del núcleo. Aunque se han propuesto varias hipótesis, no se conoce la razón de esta asimetría. Lawrence Rudnick, de la Universidad de Minnesota, sugirió que el chorro cambia periódicamente de lado. Entonces, cabe esperar que en un período largo se suministre a los lóbulos aproximadamente la misma cantidad de energía. La oscilación podría expli-



5. IMAGEN DE GRAN RESOLUCION del chorro interno. Aumenta una pequeña área cercana al centro de la galaxia. El núcleo galáctico es el óvalo rojo de la parte inferior derecha. Las áreas coloreadas hacia el nordeste del núcleo son "nudos", regiones de radioemisión particularmente intensa. Como los electrones relativistas del plasma pueden retener su alta energía sólo durante unos 50 años, deben reacelerarse para llegar al lóbulo interno, que está a 3000 años-luz del núcleo. Los nudos pueden ser regiones turbulentas del chorro, donde se reaceleran los electrones. La imagen se tomó usando el VLA en una configuración que resuelve detalles de hasta 80 años-luz de tamaño, a la distancia de Centauro-A.

carse por la interacción entre el chorro y el denso gas de la región del núcleo. Para otros investigadores los chorros se extienden a ambos lados del núcleo; a causa de efectos relativistas, aumentaría la emisión del chorro dirigido hacia nosotros, mientras que la emisión del que se aleja se vería fuertemente reducida [véase "Chorros cósmicos", de Roger D. Blandford, Mitchell C. Begelman y Martin J. Rees; INVESTIGACIÓN Y CIENCIA, julio de 1982]. Sin embargo, esta explicación requiere que la velocidad con que se mueve gran parte del plasma se acerque a la de la luz, cuestión que parece bastante improbable en Centauro-A.

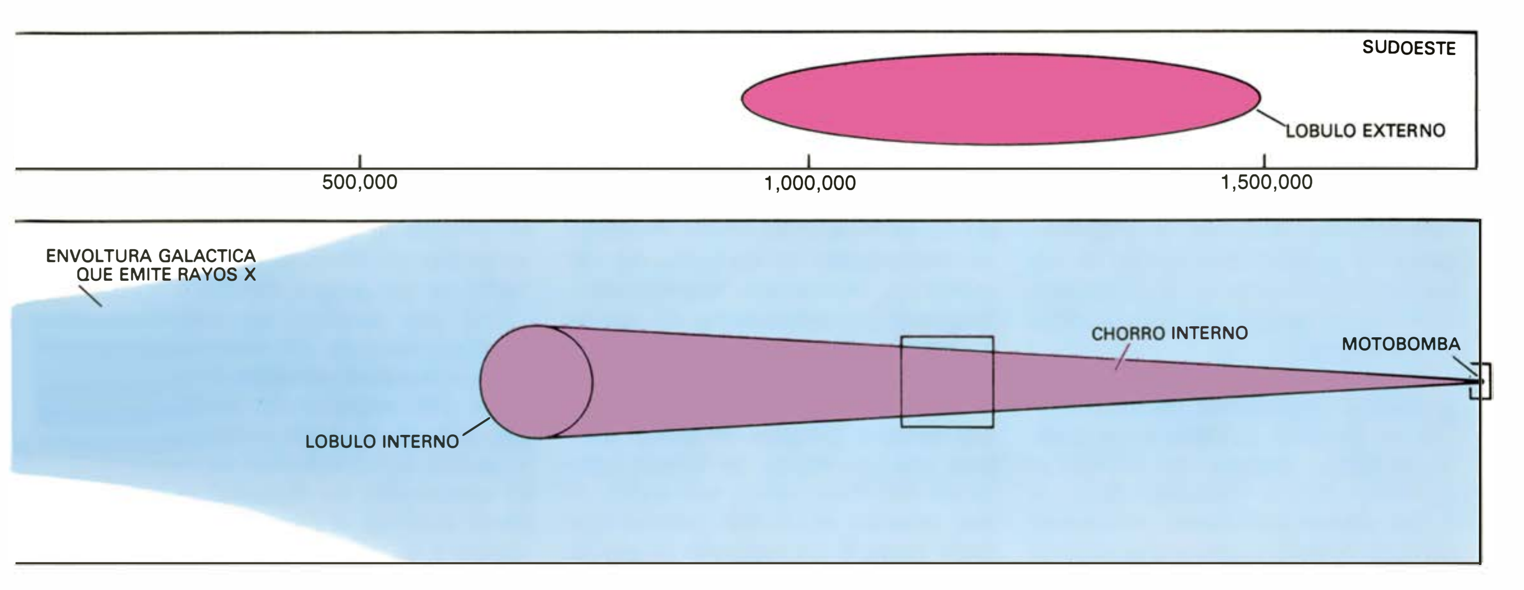
En radiointerferometría la resolución en las observaciones crece al aumentar la línea basal (es decir, la distancia entre antenas). La resolución

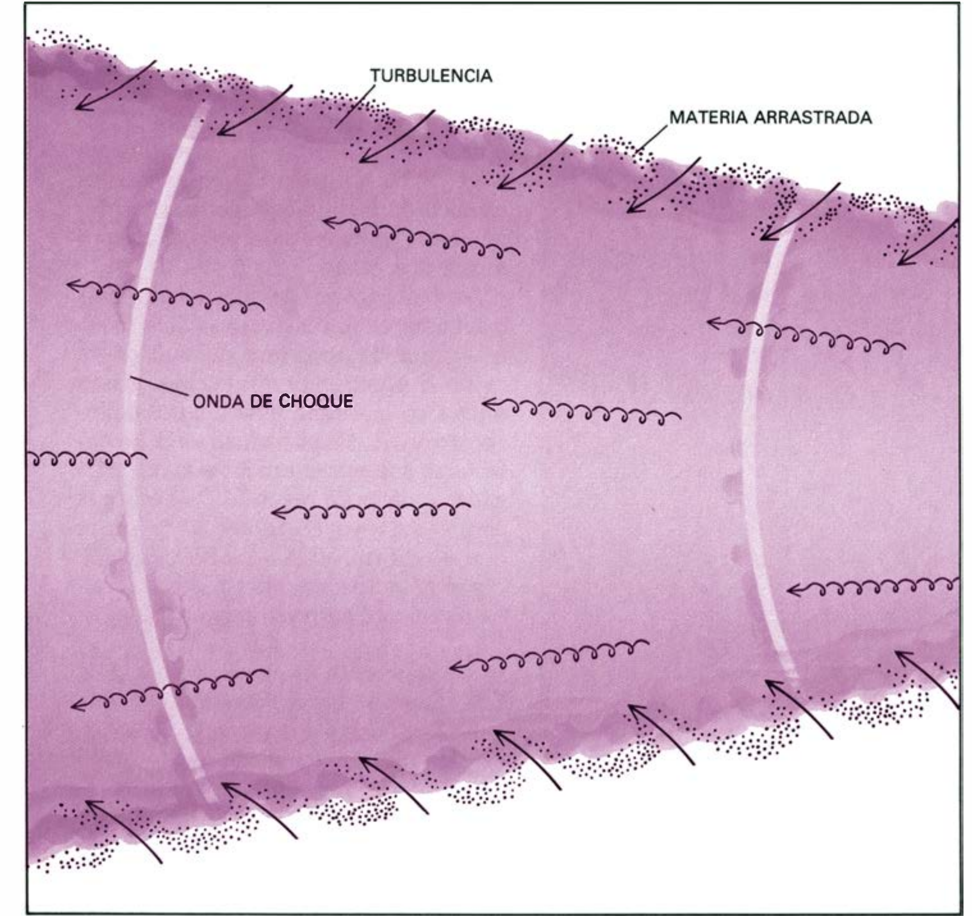


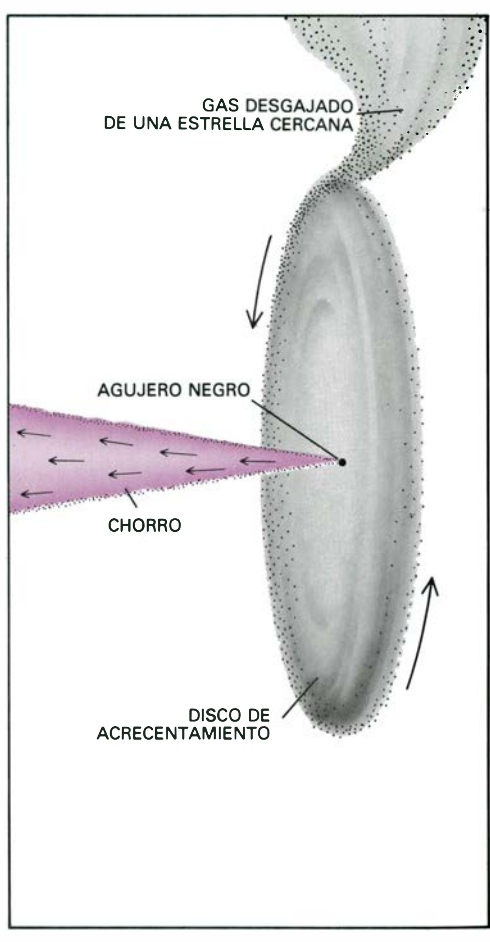
6. DINAMICA DE LA REGION DE EMISION de Centauro-A, esquematizada en recuadros con una escala progresivamente más pequeña. La extensión total de la región de emisión es de tres millones de años-luz (*recuadro superior*). Las radioemisiones están dibujadas en rojo, mientras que las de rayos X lo están en azul. La pareja de lóbulos internos, a cada lado del centro de la galaxia, son áreas de radioemisión bastante intensa, dentro de una envoltura difusa exterior que también emite. El chorro interno que conecta el núcleo con el lóbulo interno Norte emite mucha radiación en ambas zonas espectrales, rayos X y radiofrecuencias (*recuadro central*). Se sitúa dentro de una región mayor de emisión difusa de rayos X que viene del medio interestelar. La región difusa de emisión de rayos X se extiende hacia el lóbulo medio Norte. Uno de los bordes del lóbulo medio es también el lugar de aparición de un grupo de filamentos extensos y brillantes, fuentes puntuales azules de radiación visible. Una hipótesis que da cuenta de muchas de las observaciones

que se puede conseguir en un radiointerferómetro se ha perfeccionado con el despliegue de antenas en varios continentes. Se graban los datos de cada instrumento en cinta magnética, para compararlos y correlacionarlos luego con un ordenador. Robert A. Preston y sus colegas del Laboratorio de Propulsión a Chorro del Cal Tech emplearon radiotelescopios en Australia y en Sudáfrica para observar Centauro-A, llegando a precisar, en el núcleo galáctico, detalles de hasta tres milisegundos de arco.

Preston y sus colegas encontraron en el interior del núcleo de Centauro-A un chorro de una longitud de unos 50 milisegundos de arco, lo que equivale a cuatro años-luz. Este chorro pequeño emerge con igual ángulo que el gran radiochorro encontrado con el VLA.







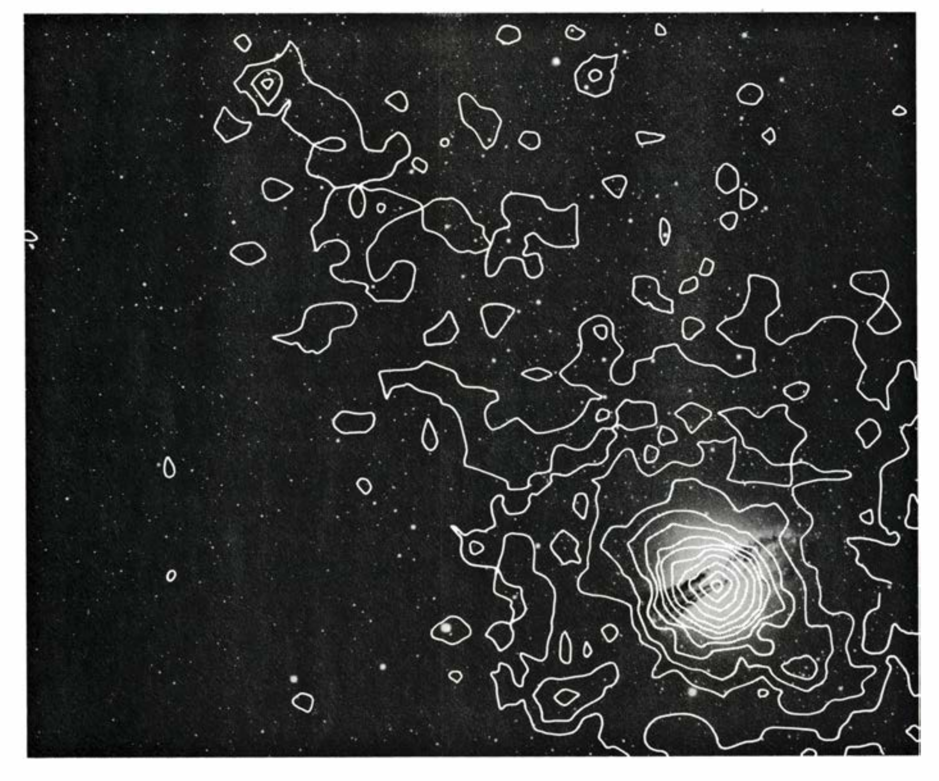
recientes de radiofrecuencias y rayos X de Centauro-A es la siguiente: la motobomba que genera el radioplasma tiene en el centro un agujero negro con una masa de mil millones de soles (*recuadro inferior*). El agujero negro está rodeado por un disco de acrecentamiento de forma toroidal, constituido por gas y polvo. El combustible para la motobomba viene del gas que cae; éste puede desgajarse, por efectos de marea, de las atmósferas de estrellas cercanas al disco. La interacción del gas con el agujero negro y el disco de acrecentamiento produce un haz de electrones y campos magnéticos estrechamente colimado: el chorro interno. A medida que el chorro se aleja del núcleo, la turbulencia creada en los bordes del mismo consigue capturar el gas exterior transportándolo con el plasma (*recuadro inferior, en el centro*). Dentro del chorro, ondas de choque van acelerando de nuevo los electrones relativistas. Cuando el plasma llega a la región del lóbulo medio, parte del gas que entró se enfría y contrae formando estrellas (*recuadro inferior, a la izquierda*).

Esta notable colimación del haz, de 4 a 20.000 años-luz, significa que la motobomba del interior núclear mantiene durante largos períodos una persistente orientación espacial. El pequeño chorro interno está muy cerca de la motobomba y es probablemente el punto inicial donde el haz de plasma empieza a ser colimado.

Para completar el radiomapa de la galaxia tenemos el núcleo. En radioobservaciones, el núcleo aparece como una fuente muy potente en el centro mismo de la galaxia; con los instrumentos actuales, no puede resolverse en sus componentes. Mientras se estudiaba mediante radiointerferometría la estructura pormenorizada de Centauro-A, se iba recogiendo una nueva y valiosa información.

Entre el instrumental del observatorio Einstein, a bordo de un satélite artificial, lanzado en noviembre de 1978, iba un telescopio de rayos X, que constituía el primer instrumento capaz de detectar y cartografiar rayos X blandos de fuentes no solares. Dentro de la zona del espectro electromagnético perteneciente a rayos X, los rayos X blandos son los dotados de energía relativa más baja: van de unos 1000 a 10.000 electronvolt. Los rayos X duros corresponden a fotones de mayor energía: 10.000 a 50.000 electronvolt. Los dos detectores principales colocados a bordo del satélite resolvían estructuras de rayos X de 1,5 minutos de arco y de cinco segundos de arco respectivamente, dando por primera vez a los astrónomos que estudian rayos X un poder de resolución del rango del que poseen los telescopios ópticos y radiotelescopios.

Poco después de que el satélite Einstein fuese lanzado al espacio, Feigelson y Schreier, junto con Riccardo Giacconi, del centro astrofísico del Observatorio Universitario de Harvard y del Observatorio Astrofísico Smithsoniano, dirigieron los instrumentos del satélite a Centauro-A; encontraron una riqueza de detalles sin precedente en la astronomía extragaláctica de rayos X. A una distancia y ángulo del núcleo similares a las que está el radiolóbulo intermedio del Norte, existe una región de una longitud de 24.000 años-luz que emite rayos X. La radiación es más intensa de lo que podría esperarse para una emisión de sincrotrón. Probablemente sea radiación térmica; su fuente podría estar en una nube de gas a una temperatura de 10 millones de grados Kelvin. El gas se calentaría al ser comprimido por la turbulencia en el radioplasma del lóbulo intermedio.



7. MAPA DEL CONTORNO DE RAYOS X, conseguido en el Observatorio Einstein, superpuesto a una imágen óptica de NGC 5128. El cuerpo principal de la galaxia está en la parte inferior derecha. Los filamentos visibles y estrellas jóvenes que se hallan cerca del lóbulo medio Norte en la parte superior izquierda están situadas a 130.000 años-luz del núcleo. Las emisiones de rayos X fuera del cuerpo principal de la galaxia coinciden con los filamentos visibles. La imagen óptica se obtuvo por medio de la línea de emisión que resulta de la recombinación de un electrón y un protón para formar un átomo de hidrógeno.

La galaxia se encuentra rodeada por una envoltura que emite rayos X menos intensos, extendiéndose hasta una distancia de 10.000 años-luz del núcleo. Probablemente, la emisión difusa se genera por una nube de gas interestelar con una temperatura de unos 20 millones de grados Kelvin y una masa total de 200 millones de veces la solar. Con la detección de gas caliente mediante los instrumentos a bordo del Observatorio Einstein, tenemos una de las pocas veces en que un medio interestelar ha sido observado dentro del cuerpo visible de una galaxia elíptica.

Con una densidad aproximada de una partícula cada 300 centímetros cúbicos, el medio interestelar descubierto en las observaciones del satélite Einstein es bastante denso, comparado con el que nos encontramos corrientemente en astronomía. La densidad existente puede confinar el plasma en el radiochorro y en el radiolóbulo interno y contribuir, por consiguiente, al mantenimiento de su forma. Tal confinamiento podría resultar crítico. Sin él, el radiochorro podría perder su colimación y los radiolóbulos expandirse, diluyéndose su densidad de energía. Si las hipótesis son correctas, la estructura de NGC 5128 confirma el confinamiento térmico de las radioestructuras, un fenómeno que había sido predicho, pero nunca observado.

Sin embargo, el descubrimiento más excitante realizado con el satélite Einstein fue el de un chorro de emisión de rayos X procedente del núcleo, más o menos en la misma posición del radiochorro VLA. Fue la primera vez en detectarse una estructura intensa tan alargada en datos de rayos X. De hecho, el chorro de rayos X se descubrió antes que el chorro VLA; los hallazgos en rayos X motivaron gran parte del trabajo reciente en radiofrecuencias.

No solamente el chorro de rayos X tiene el mismo tamaño y forma que el radiochorro, sino que, además, existen "nudos" en las mismas posiciones: regiones pequeñas donde las emisiones alcanzan una notoria intensidad. Las semejanzas estructurales, junto con el espectro continuo de emisiones y sus propiedades de polarización, nos llevan a concluir que la emisión en rayos X de Centauro-A es también radiación de sincrotrón.

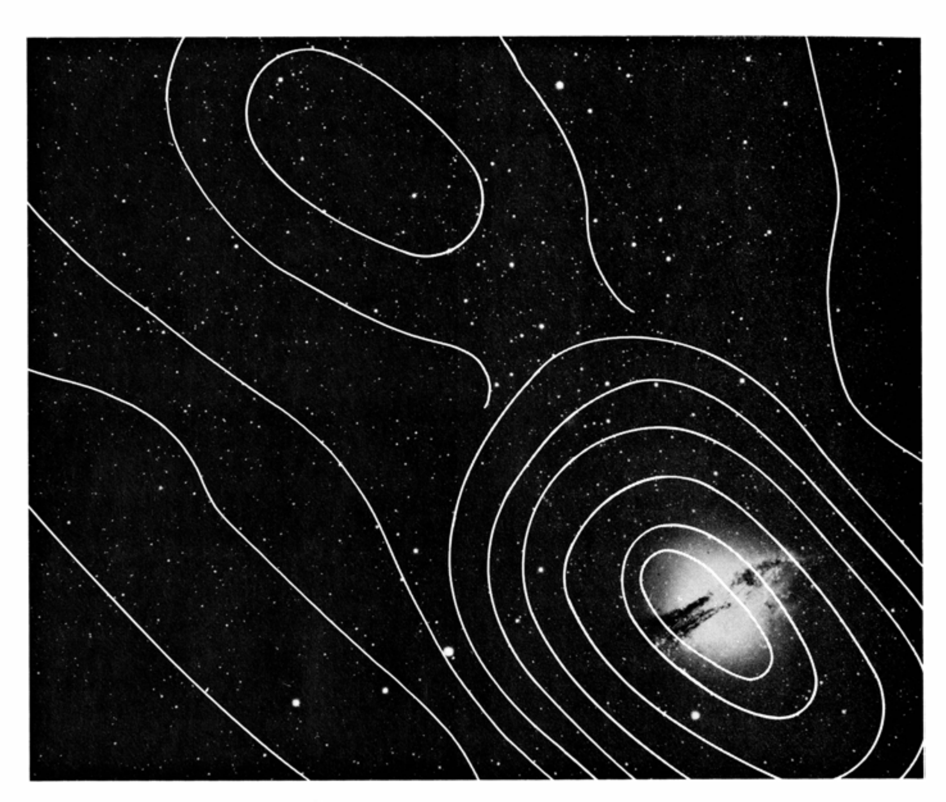
Cuando se analizaron los datos tras más de dos años de funcionamiento del

satélite Einstein, se descubrió sólo otro chorro galáctico de rayos X como el de NGC 5128. Este segundo chorro está en la siguiente galaxia activada más cercana: la M 87, galaxia que dista de nosotros tres veces lo que nos separa de Centauro-A. No obstante, los instrumentos del satélite carecen de resolución suficiente para mostrar detalles estructurales del chorro de M 87. El chorro de rayos X de Centauro-A tiene rasgos que pueden distinguirse en sí mismos, lo que le convierte en un objeto astronómico único.

¿Qué nos dice el mapa detallado de NGC 5128, pacientemente agrupado en rayos X y radiofrecuencias, acerca de los procesos físicos que se desarrollan en la galaxia? Consideremos los chorros internos, el de rayos X y el de radioondas. Por lo que se refiere a la observación astronómica, la componente más significativa del chorro es el flujo de electrones relativistas (electrones que se mueven a una velocidad cercana a la de la luz). Estos electrones son la fuente de la radiación de sincrotrón. Sin embargo, no se deduce que, a causa de que los electrones se muevan a velocidades relativistas, el chorro en conjunto tenga tal velocidad. El haz contiene, además de los electrones, un volumen considerable de gas, con una velocidad mucho más baja que la de los electrones. La mejor estimación actual sobre la velocidad global de flujo del chorro se cifra en unos 5000 kilómetros por segundo.

La conclusión de que los fotones de rayos X detectados por el satélite Einstein surgen de un proceso de sincrotrón impone algunas severas restricciones al modelo del comportamiento de chorro. Los fotones de rayos X tienen una energía de unos 2000 electronvolt. En un proceso de sincrotrón, los fotones son emitidos por electrones relativistas con una energía de 3×10^{13} electronvolt. Electrones con tamaña energía no permanecen ligados a un campo magnético mucho tiempo sin emitir fotones.

Si continúa la emisión, el electrón terminará por agotar su energía y no podrá emitir ningún fotón más. Desde luego, dadas las energías de los electrones y de los fotones de rayos X, el agotamiento de energía ocurriría a lo sumo en 50 años. Sin embargo, el chorro de rayos X y radioondas de Centauro-A llega hasta el lóbulo interno, que está a 20.000 años-luz del núcleo. Por tanto, si los electrones viajan a una velocidad muy cercana a la de la luz, tendrán que reacelerarse muchas veces para conseguir llegar al lóbulo interno.



8. MAPA DEL RADIOCONTORNO, obtenido con el radiotelescopio de 64 metros del Observatorio de Radioastronomía de Australia, superpuesto a la misma imagen óptica de NGC 5128 presentada en la ilustración precedente. Como en el caso de las emisiones de rayos X la región de radioemisión más intensa fuera del cuerpo principal de la galaxia (el lóbulo medio Norte) coincide con los filamentos visibles y las estrellas jóvenes. El solapamiento de las fuentes de radiación en las tres frecuencias (a saber: visible, radio y rayos X) apoya la hipótesis según la cual habría un mismo proceso responsable para todas ellas.

Los electrones relativistas del haz podrían acelerarse por alguno de los tres mecanismos siguientes que detallamos. Supone uno la existencia, en el chorro, de ondas de choque parecidas a las que se forman en la atmósfera terrestre cuando un avión vuela a una velocidad mayor que la del sonido. Tales choques podrían resultar de la colisión entre el chorro y las nubes interestelares; también podrían formarse internamente, si el haz fuera no estacionario o inestable. Los nudos del haz corresponderían a regiones donde el plasma está sujeto a fuertes choques.

En el segundo mecanismo, es la energía de la turbulencia interna del haz la que acelera a los electrones. Pueden formarse turbulencias e inestabilidades en una capa fronteriza entre el haz del plasma y el medio que lo rodea. La turbulencia se propagaría hacia el interior del haz en forma de ondas, cuyo movimiento se vería amortiguado por la interacción con los electrones. En la interacción con estas ondas, los electrones serían acelerados. El resultado final del proceso sería una transferencia de la energía de flujo de plasma a energía cinética de las ondas y después en aceleración de los electrones relativistas.

De acuerdo con la tercera posibilidad, los electrones aumentan su energía de nuevo mediante colisiones con los protones del haz. Los protones del plasma no pueden detectarse directamente puesto que no emiten la suficiente cantidad de radiación. Sin embargo, por la misma razón, los protones pueden retener su alta energía cinética un tiempo más largo que los electrones. Estas colisiones terminarían por agotar la reserva de energía protónica, pero el proceso sería lento y podría tratarse de una reserva suficiente para mantener el plasma hasta llegar al radiolóbulo interno.

No se conoce bien en virtud de qué proceso el haz de protones, electrones y campos magnéticos pasa del lóbulo interno al medio. Funcione del modo que sea, cuando el plasma llega al lóbulo medio podría tener un importante papel en la formación de estrellas jóvenes y en las regiones de líneas de emisión observadas a una distancia de unos 130.000 años-luz del núcleo.

David S. De Young, del Observatorio Nacional de Kitt Peak, propone el siguiente escenario para la interacción de las radioestructuras con las estructuras ópticas. Los remolinos de la

capa turbulenta de los bordes del chorro pueden introducir material cercano, formado principalmente por gas y polvo; este material se calienta al penetrar, por choques con el haz de plasma a una temperatura de unos 10 millones de grados Kelvin.

Si el gas que rodea al chorro tiene la misma abundancia de elementos que el Sol, lo cual es una hipótesis razonable para una galaxia adulta, se enfriará a unos 10.000 grados Kelvin en un tiempo que va de 10 a 100 millones de años. A la temperatura de 10.000 grados, el gas mostraría líneas de emisión en el visible, como las que se ven en el cuadrante nordeste de NGC 5128. Es más, en el tiempo que el gas tardaría en enfriarse habría podido viajar una distancia entre 33.000 y 330.000 años-luz, lo que lo situaría cerca de la posición donde se ve la emisión óptica en NGC 5128 y en otras radiogalaxias diferentes, donde las líneas de emisión en el visible coinciden con los radiolóbulos.

Una fracción del gas introducido continuaría enfriándose; con el tiempo llegaría a temperaturas bastante más bajas que 10.000 grados. El gas comenzaría entonces a colapsar por atracción gravitatoria, pudiendo formarse estrellas en su seno. De Young mantiene que los cúmulos con emisión en el visible observados en el cuadrante nordeste de NGC 5128 apoyan la conclusión de que las estrellas se están formando en la región del lóbulo medio.

El plasma contribuye a la formación de estrellas aportando materia del medio; a su vez, las estrellas pueden contribuir a mantener la energía del haz de plasma a medida que se mueve desde el centro de la galaxia hacia el exterior. Algunas de las estrellas que se forman en el cuadrante nordeste pueden tener masas superiores a 10 masas solares. Estas estrellas tan masivas tienen una vida bastante corta: aproximadamente unos 10 millones de años. Al final de su relativamente corta vida estas estrellas masivas estallarían como supernovas.

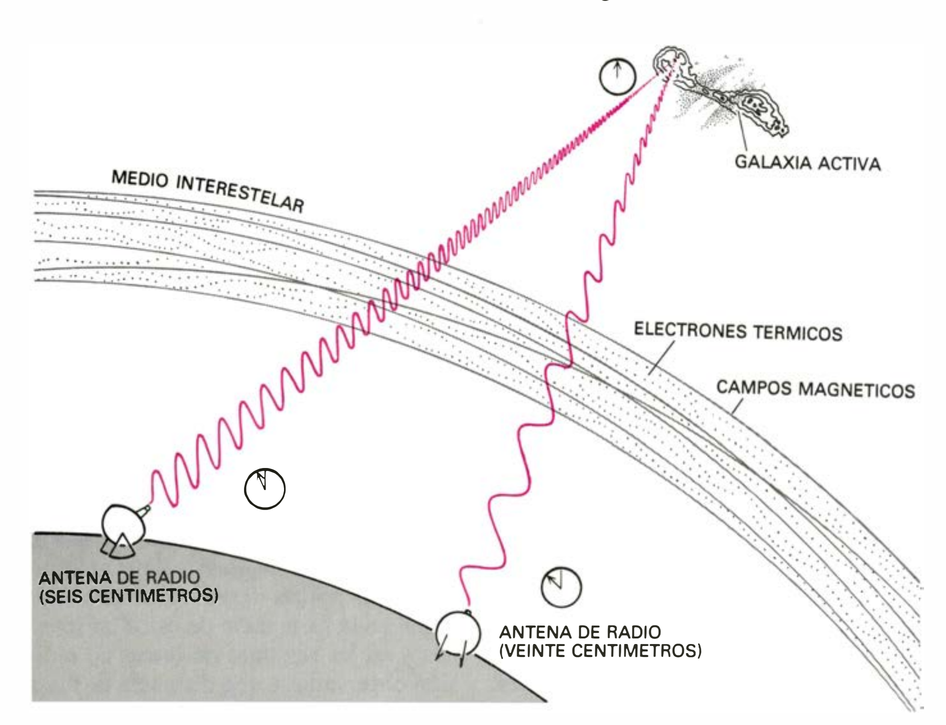
El efecto de estas explosiones podría ser el de devolver la energía perdida a los electrones del haz de plasma, por lo que el flujo en el cuadrante nordeste puede automantenerse: el haz de plasma da origen a la formación de estrellas gigantes que al morir irán estallando y devolviendo su energía al haz.

Esta descripción de cómo el haz de plasma puede adquirir energía y de cómo se forman los filamentos visibles se basa en abundantes datos experimentales. Dos de las cuestiones más fundamentales acerca de Centauro-A y de otras galaxias activas deben resolverse, en parte, mediante conjeturas. Estas cuestiones son: ¿cuál es el origen de la actividad galactica? ¿Cómo funciona la motobomba central?

Es posible que la actividad observada en NGC 5128 esté direcamente relacionada con la peculiar estructura de la galaxia, en la cual se combinan una elipse y un disco. Hace unos treinta años, Walter Baade y Rudolph Minkowski, de los observatorios de Monte Wilson y Monte Palomar, propusieron que la forma de NGC 5128 resultaba de una fusión entre una galaxia espiral y otra elíptica. Un trabajo teórico reciente, realizado por Allan D. Tubbs, del Observatorio Nacional de Radioastronomía, demostró que el gas y polvo desgajado de alguna galaxia cercana puede precipitarse sobre una galaxia elíptica y depositarse en ella en una disposición que recuerda mucho la de la franja de polvo que se observa en NGC 5128.

Tal proceso de fusión ha sido denominado canibalismo galáctico. Este canibalismo se utiliza para explicar intensas radiofuentes, distintas de Centauro-A. Cierto tipo de cúmulos de galaxias que tienen en su centro una galaxia gigante son radiofuentes cinco veces más intensas que el promedio normal. Se ha afirmado que la galaxia gigante central se originó por repetidas colisiones gravitatorias en lo más profundo del cúmulo. Estas colisiones generan una fuerza de fricción que hace que las galaxias mayores adquieran un movimiento espiral lento que les aproxima hacia el centro del cúmulo. Durante el tiempo de vida del cúmulo, las galaxias más masivas continúan colapsando, formando, en último término, una única unidad central muy masiva. La galaxia central gigante engulle a las galaxias cercanas más pequeñas.

El origen de la actividad en radiofrecuencias de estas galaxias gigantes radica, casi con toda seguridad, en el material que se acumula durante este proceso de canibalismo. Este aumento de gas, polvo y estrellas proporciona el combustible para la motobomba del centro galáctico. Un canibalismo más limitado en NGC 5128 podría también haber desencadenado violenta actividad en el núcleo galáctico y, por tanto, emisión en radiofrecuencias. Una objeción a la hipótesis del canibalismo es que Centauro-A está aislada: sus únicas compañeras, en un radio de un millón de años-luz, son unas pocas galaxias enanas. Podría ser que la galaxia hubiese engullido a todas sus vecinas; las emisiones no térmicas en radiofrecuencias y rayos X serían signos de una indi-



9. EFECTO FARADAY, que gira el plano de polarización de la radiación que atraviesa nuestra galaxia. La radiación proveniente de una galaxia activa, así Centauro-A, está generalmente polarizada en un plano, ya que se produce en procesos de sincrotrón. Al pasar la radiación a través del medio interestelar de nuestra galaxia, interacciona con electrones térmicos y campos magnéticos; en virtud de lo cual, gira el plano de polarización de las ondas. El ángulo de rotación es proporcional a la longitud de onda: a mayor longitud de onda, mayor ángulo de giro. Las radioobservaciones de Centauro-A que se hicieron a principios de los años sesenta proporcionaron la primera confirmación experimental del efecto Faraday en la Vía Láctea. Las observaciones han apoyado también la conclusión según la cual el espectro de banda ancha emitido por la galaxia activa Centauro-A tiene como origen un proceso de sincrotrón.

gestión cósmica a consecuencia de una comida opulenta.

¿Cómo puede utilizar la motobomba central el combustible que le proporciona este proceso de canibalismo? El combustible se consume, tal vez, de un modo más rentable si es en forma de nube de gas. Observando con el VLA, Jan M. van der Hulst, del Observatorio Holandés de Westerbork, acaba de descubrir nubes de gas de hidrógeno neutro a una distancia de 500 años-luz del núcleo de Centauro-A. El comportamiento del movimiento de las nubes sugiere que están cayendo hacia el núcleo. Dichas nubes podrían proporcionar cada año a la motobomba el combustible equivalente como mínimo a una décima parte de la masa solar. Si suponemos que la motobomba tiene un rendimiento del 10 por ciento en el proceso de convertir masa en energía, una décima parte de la masa solar por año podría ser más que suficiente para dar cuenta de la energía producida y, concretamente, de la extensión de la radioregión.

El combustible podría proceder, asimismo, de las atmósferas de las estrellas situadas en la región central de la galaxia. Jack G. Hills, del Laboratorio Nacional de Los Alamos, entre otros, sugirió que las estrellas inmediatas al núcleo estarían sometidas a grandes fuerzas de marea, de manera que la intensidad del campo gravitatorio variaría considerablemente de un lado a otro de la estrella. En virtud de lo cual la estrella se fragmentaría o, cuando menos, perdería su atmósfera.

La materia es, pues, un combustible aprovechable para la motobomba del núcleo de la galaxia. La cuestión siguiente estriba en dilucidar de qué manera las nubes de gas, que tienden a rodear el núcleo, caen capturadas por éste. Cabría la posibilidad de que se debiera a la colisión de dos nubes de gas en las proximidades del núcleo. En tal colisión una nube podría perder momento angular y precipitarse sobre el núcleo, mientras que la otra, que habría aumentado su momento angular, se alejaría del mismo.

Una vez resuelto el problema del combustible y sugerido un mecanismo alimentador por cuya virtud el combustible puede depositarse en la motobomba, ha de abordarse qué ocurre en la propia motobomba. Para una tal consideración nos sería muy útil poder estimar el tamaño de la máquina. Para calcular el tamaño de una fuente cualquiera de radiación se mide el período de tiempo para el cual la intensidad de la emisión varía. El razonamiento en que se basa esta estimación es el siguiente. Si existen en una fuente dos regiones extensas separadas que emiten y dejan de emitir simultáneamente, deben tener algún tipo de comunicación: una señal física debe propagarse entre ellas. Puesto que lo más rápido que pueden propagarse las señales es a la velocidad de la luz, podemos concluir que si la emisión varía de modo significativo, digamos sobre dos horas, la propia fuente no superará las dos horas-luz en esa misma dirección.

A pesar de que este argumento no es incontrovertible, resulta probablemente válido para una fuente con una superficie de emisión continua; lo hemos usado para estimar el tamaño de la motobomba central de NGC 5128. Las observaciones sobre la variabilidad del radioflujo y del flujo de rayos X procedentes del núcleo de Centauro-A muestran que la emisión varía significativamente en un intervalo de meses. Se observan, sin embargo, variaciones más rápidas, llegando a ser menores de 24 horas. Por tanto, puede concluirse que el núcleo tiene partes diferenciadas cuyo tamaño oscila desde meses-luz hasta días-luz.

Aunque cualquier explicación que se proponga de lo que ocurre en las subestructuras de la motobomba tiene que apoyarse en la especulación incontrastada, un punto parece claro. Cualquiera que sea el funcionamiento de la motobomba, no se trata de la nucleosíntesis, proceso de fusión nuclear que proporciona la energía de la mayoría de las estrellas. La hipótesis del diez por ciento en el rendimiento de la motobomba central descarta por completo la nucleosíntesis, cuya eficiencia no llega al uno por ciento y requiere, en ese supuesto, más combustible y más tiempo para generar la estructura en radiofrecuencia de Centauro-A que los que parecen desprenderse de las observaciones más recientes. Es más, la nucleosíntesis produce radiación térmica, en vez de la emisión de sincrotrón no térmica que se registra en NGC 5128.

Un modo más eficiente de extraer energía de la materia se produce cuando ésta colapsa en un campo gravitatorio intenso. En el caso en que esto ocurre, la materia gana una cantidad considerable de energía cinética. Si este material muy energético colisiona con otras estructuras del núcleo, la energía se liberará en forma de ondas electromagnéticas de gran energía e, incluso, como partículas a gran velocidad.

Para que éste fuese el funcionamiento de la motobomba central de Centauro-A tendría que existir alguna fuente que produjese un campo gravitatorio muy intenso en el centro mismo de la galaxia. Tal campo podría obtenerse de un objeto totalmente colapsado: un agujero negro con una masa aproximada de 1000 millones de veces la solar. Si tal objeto existiese en el centro mismo de la galaxia debería sin duda estar girando; es difícil imaginar un proceso que conduzca a la creación de un agujero negro sin comunicarle, al mismo tiempo, un momento angular considerable.

Este cuerpo colapsado, en rotación, produciría tres efectos principales. En primer lugar, daría origen al pozo de potencial donde puede precipitarse el combustible. Segundo, la motobomba se orientaría según su eje de giro. La rotación produciría, en el material circundante, un movimiento de precesión en torno a la motobomba y la formación de un disco de gas, el llamado disco de acreción. Por último, el agujero negro poseería un campo magnético asociado a él, cuyas líneas de fuerza podrían acelerar partículas cargadas hasta altas energías, extrayendo progresivamente energía del agujero negro.

Imagínese una nube de gas relativamente fría precipitándose sobre un disco de acreción (o acrecentamiento). El disco podría presentar una forma toroidal con el agujero negro en su centro. Parte del gas que caería sería colimado en el interior de los reducidos límites del agujero negro y expelido hacia afuera a lo largo del eje de rotación del agujero, gracias a la presión de radiación de la emisión procedente de la superficie interna del disco de acreción. De este modo, el disco podría colimar y acelerar el haz de partículas que forma el radiochorro interno.

Por el momento, esto es lo más lejos a que pueden llevarnos las especulaciones más audaces sobre las galaxias activas del tipo de Centauro-A. Galaxia que continúa polarizando el interés de la investigación científica. Debido a su proximidad y a su enorme valor intrínseco seguirá atrayendo la atención de muchos en los próximos años. El trabajo en curso ha acumulado gran cantidad de información sobre la estructura y los procesos físicos que se desarrollan en su seno. El aspecto más obscuro de una galaxia activa es, sin embargo, el mecanismo en que se basa su enorme región de emisión. Confiamos en que el estudio de Centauro-A ayudará a la resolución de este problema, fundamental en astrofísica.

Ciencia y sociedad

El metro

Se ha considerado siempre el metro como la unidad más significativa en los diversos sistemas que se emplean en la ciencia, la técnica, el comercio y en múltiples aspectos de nuestra vida diaria.

Porque medir es resolver, avanzar, desde que la revolución del Neolítico hizo posible el sedentarismo, la agricultura, los excedentes y el intercambio. Para medir lo que se cambiaba fueron surgiendo unidades necesariamente relacionadas con la propia experiencia humana. El hombre comparaba el exterior consigo mismo y medía el terreno a pasos, la longitud a brazas, el peso en cargas, los intervalos de tiempo en latidos de su corazón.

El desarrollo de las ciudades (concentración de servicios y de intercambios) exigía fijar y controlar las unidades. En las islámicas, sin plazas, la vida pública se teje frente a las puertas, en el mercado a extramuros; y en esas puertas aparecen tallados pies y varas (Zafra) o pesos colgando (Granada). En la España islámica se trató de mantener la medida que establece el Corán y relacionar el codo con la medida de capacidad, eligiendo como medida de peso también la arroba. Por contra, en las ciudades cristianas, donde la vida y el mercado se arropan a la sombra de la Iglesia, será en la parte exterior de su ábside donde queden indeleblemente marcadas esas unidades, que suelen variar no sólo de región en región sino también de ciudad en ciudad. Para evitar la confusión, los gobernantes se preocuparon por unificar las unidades, sobre todo las de longitud y peso. En unos casos el pie o la vara fueron los que correspondían al rey. En otros, se cogieron unos cuantos hombres a la salida de la iglesia y se halló la media de sus pies.

Será la Revolución francesa quien inicie la unificación y racionalización de todas las unidades existentes, aparte de introducir los múltiplos decimales en el llamado sistema métrico decimal, basado en el metro geodésico. Inicialmente se trató de fijar cada unidad base por un prototipo material que la representara permanentemente. Así se hizo con el metro, el kilogramo, el ohmio, etcétera. Después, se descubrió que no estaba asegurada la invariabilidad de estas unidades materiales y que, accidentalmente, podían quedar destruidas. Se buscaron sustitutos en recetas de laboratorio que pudieran reproducirse con cierta facilidad. El progreso de la industria y del nivel de vida exigían cada vez menor imprecisión en las medidas, y por tanto en las unidades, y mayor fiabilidad en la transferencia y comparación de patrones entre los distintos centros. Un ejemplo clásico nos lo ofrece la industria de precisión, que hacia 1800 se conformaba con 1/4 de milímetro, en 1900 exigía 0,01 milímetros, en 1950 pasaba de 1/4 de micrometro, en la década de los 70 llegó a 0,01 micrometros, y hoy exige aún más.

Las distintas uniones científicas se han interesado por estos problemas creando en su seno comisiones de unidades y nomenclatura. También la Organización Internacional de Normalización tiene una comisión, la CT12, a ello dedicada. Los laboratorios nacionales de metrología se ocupan del problema. Todas estas sugerencias acaban siendo canalizadas hacia el Comité Internacional de Pesas y Medidas (CIPM), que tiene una serie de comités consultivos verticales (uno por cada unidad de base, aproximadamente) y un comité consultivo de unidades, que cubre todos los campos y prepara las proposiciones que el CIPM somete a la Conferencia General de Pesas y Medidas, cuyos acuerdos gozan de carácter legal en todos los países adheridos.

Así se ha ido elaborando el llamado sistema internacional de unidades, abreviadamente SI. Adviértase que este sistema no pretende ser una construcción intelectual que persiga, sobre todo, la perfección racional; se trata de un acuerdo convencional, un consenso puramente pragmático, que se quiere universal, preciso, cómodo y sencillo, sin olvidar su posible elegancia.

Distinguimos entre unidades básicas y unidades derivadas. La definición de toda unidad básica (metro, segundo, ampère, etc.) es la descripción de una situación experimental, que permanece invariable frente a cualquier reajuste en las otras unidades básicas, aun cuando las utilice para fijar el valor de las magnitudes de influencia o para especificar la propia situación experimental. Pero esas otras unidades, como las palabras, se utilizan sólo como símbolos y si esas unidades cambian, como si cambiara el sentido de una palabra, deberá modificarse adecuadamente la redacción de la definición para que siga describiendo la misma situación experimental. Con la excepción del kilogramo, cada unidad básica está ligada, en su realización, con una característica física: la propagación de la luz, un período, la temperatura termodinámica del punto triple del agua, el número de átomos de una cantidad de sustancia, la permeabilidad en el vacío y la intensidad luminosa de una cierta energía. El kilogramo está, por el momento, ligado a un patrón material.

El metro, que dio nombre al sistema métrico decimal y del que se derivarían el litro y el kilogramo, fue un producto de la Revolución francesa. Lo estableció en 1795 la Asamblea, siguiendo recomendaciones de la Academia de Ciencias. El metro era la diezmillonésima parte del cuadrante del meridiano terrestre según las mediciones que habían hecho en 1731 Jorge Juan y Ulloa en Sudamérica y que repitieron los científicos franceses en 1792 entre Dunkerque y Barcelona. En 1799 se repitieron las medidas y se materializó el metro en una regla de platino depositada en Francia. Al objeto de dar carácter internacional a estas medidas se convocó en 1889 la primera Conferencia General de Pesas y Medidas, a la que asistieron una serie de estados cuya lista ha ido aumentando. Se decidió entonces prescindir de la definición geodésica y fijar el metro por el famoso prototipo de platino iridiado que quedó celosamente guardado, bajo tres llaves, en Sèvres, del cual se repartieron una serie de copias de las que hay una en España.

Las crecientes necesidades de la industria sobrepasaron la precisión de este metro material. Todos los labora-

CINEMATICA	m					
DINAMICA	m	s				
MECANICA	m	s	kg			
ELECTRICIDAD	m	s	kg	A		
TERMODINAMICA	m	s	kg	A	K	
QUIMICA FISICA	m	s	kg	A	K	mol

1. Introducción de las sucesivas unidades básicas

AÑO	ORGANISMO	DEFINICION
1795	Asamblea Francesa	1/10.000.000 del cuadrante del meridiano terrestre.
1799	Asamblea Francesa	Materialización del valor anterior en una regla, a extremos, de platino, depositada en los Archivos de Francia.
1889	1.ª Conferencia General de P y M	Patrón material internacional de platino iridiado, a trazos, depositado en el BIPM. Es llamado metro internacional. (Incertidumbre 1 × 10^{-6}.)
1960	11.ª Conferencia General de P y M	1.650.763,73 λ en el vacío de la radiación del Kriptón 86 (transición entre los niveles $2p_{10}$ y 5 d_5. (Incertidumbre 1 × 10^{-8}.)
1983	17.ª Conferencia General de P y M	Longitud del trayecto recorrido en el vacío por la luz durante 1/299,792,458 segundos. (Incertidumbre 4 × 10^{-10}.)

2. Historia de las distintas definiciones del metro, cada vez más precisas

MOLECULA ABSORBENTE	TRANSICION	COMPONENTE	f EN MH_2	APLICADO A LASER	ESTABILIZADO CON CELULA DE	INCERTIDUMBRE
CH_4	ν_3, P (7)	F_2	88376181,6	He-Ne	METANO	1,3 × 10^{-10}
$^{127}I_2$	17-1, P (62)	o	520206808,5	He-Ne COLOREADO	IODO CON PUNTO FRIO A 6° C	6 × 10^{-10}
$^{127}I_2$	11-5, R (127)	i	473612214,8	He-Ne	IODO CON PUNTO FRIO A 15° C	1 × 10^{-9}
$^{127}I_2$	9-2, R (47)	o	489880355,1	He-Ne	IODO CON PUNTO FRIO A −5° C	1,1 × 10^{-9}
$^{127}I_2$	43-D, P (13)	a_3	582490603,6	Ar^+	IODO CON PUNTO FRIO A −5° C	1,3 × 10^{-9}

ATOMO 86 Kr	$2p_{10}$ a $5d_5$		λ = 605780210,2 fm			4 × 10^{-9}

*3. Realización práctica de la definición del metro mediante láser estabilizado por absorción saturada (**arriba**) y mediante lámparas espectrales (**abajo**)*

torios de metrología se ocuparon de buscar una experiencia, un montaje de laboratorio, que permitiera reproducir a voluntad y con mayor precisión la longitud metro. Finalmente, en 1960 la undécima Conferencia General de Pesos y Medidas aprobó una resolución de largo alcance: el metro se obtenía por la suma de un cierto número (muy elevado) de longitudes de onda de una radiación naranja emitida por un átomo de kriptón 86, reduciéndose la incertidumbre de la medida a 10^{-8}. A este cambio siguió la sustitución en 1968 del segundo "astronómico" por el segundo "físico", basado en una frecuencia del cesio 133 y alcanzando una incertidumbre de 10^{-13}.

Resultaba paradójico e incómodo que el metro tuviera una precisión peor que el segundo, lo que redundaba en que la exactitud en las medidas de la velocidad de la luz quedara limitada por la precisión del metro. Por otro lado, las exigencias de los cálculos de astronáutica y astronomía exigían manejar con seguridad la constante c, la velocidad de la luz que aparece en muchos cálculos. Todo ello llevó a la 15.ª C.G.P. y M. de 1975 a tomar una resolución histórica: recomendar que la velocidad de la luz, es decir, de las radiaciones electromagnéticas, dejara de ser una constante física a medir y se convirtiera en una constante de paso, con un valor fijo; justamente c = 299792458 metros por segundo.

Este acuerdo y el enorme progreso obtenido con láseres estabilizados hirió de muerte a la definición de 1960 para el metro, cuya realización podría disfrutar de incertidumbres más próximas a la del segundo, sin más que referirla a él. Si en un segundo la luz recorre c metros, recorrerá 1 metro en $1/c$ segundos. Pero hasta encontrar ensayos experimentales fiables los laboratorios de metrología hubieron de trabajar duro, ya que la definición, sin un procedimiento, ya probado, para repetir a voluntad el metro no tenía valor. Gracias al esfuerzo de diversos grupos de investigación en varios países y al cotejo continuo de sus resultados pueden hoy ofrecerse varios caminos experimentales para realizar "en casa" el metro según la nueva definición. Aun cuando la incertidumbre garantizada es sólo de 10^{-10}, es decir, de la décima cifra decimal, este límite está ya sobrepasado con lo que el metro puede codearse en precisión con el segundo.

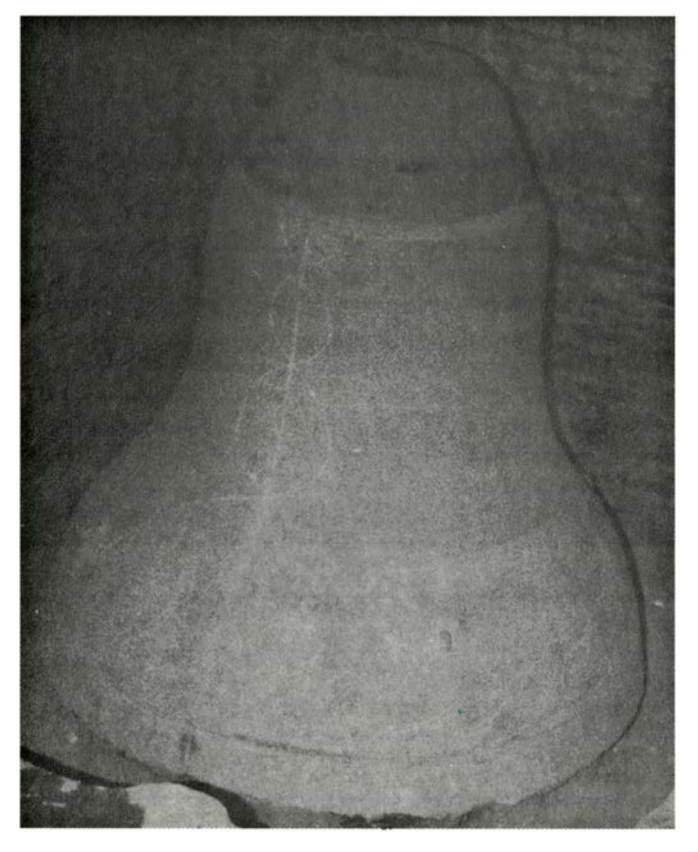
La nueva definición es más sencilla y clara, deja gran libertad en el método experimental y no introduce variación alguna en el metro, reduciendo la incertidumbre en su realización experimental. Puede seguirse usando una longitud de onda dada, que ahora se obtiene midiendo su frecuencia en segundos y usando la fórmula $\lambda = c/f$, donde lambda (λ) designa la longitud de onda y f la frecuencia. En el caso de usar kriptón se gana una cifra decimal más.

Pero también pueden usarse las radiaciones de una lista de láseres estabilizados por absorción saturada cuyas frecuencias se han determinado ya, por ciertos laboratorios metrológicos, con incertidumbres de hasta 10^{-10} y que cada día ganan mayor precisión.

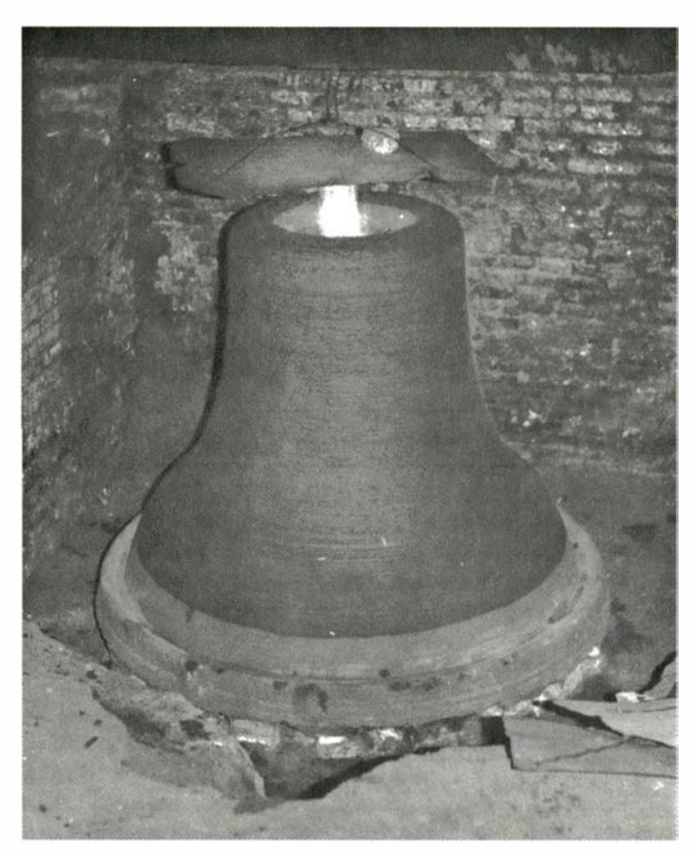
Y llegamos así a octubre de 1983, en que la XVII C.G. P. y M. decide: *1.* El metro es la longitud del trayecto recorrido en el vacío por la luz durante 1/299 792 458 segundos; *2.* La definición de metro, en vigor desde 1960, fundada en la transición entre los niveles $2p_{10}$ y $5d_5$ del átomo de kriptón 86 queda derogada. (Leonardo Villena.)

Fundición medieval

Lo mismo que el inventario notarial de los bienes de una persona nos puede dar una idea de su situación económica y la relación de los libros de su biblioteca de su nivel cultural, las anotaciones de los gastos que se han producido en la realización de una obra o de una operación comercial, los libros de cuentas, pueden ilustrar sobre la mar-



1. Molde completo de la campana



2. Cocido del macho

cha de la misma y aportar detalles relacionados con ellas. Viene esto a cuento de un cuadernillo donde se anotaron los gastos que se produjeron con motivo de la fundición de una campana en 1405. Durante siglos, traspapelado entre libros de registros de bautizos, bodas y entierros, han permanecido en el archivo de la parroquia de S. Miguel de Burjasot (Valencia, España) ocho hojas de papel, dobladas y cosidas por su mitad, en donde, cuidadosamente, día a día, se escribieron todos los pagos que hubo que hacer para fundir una campana, cuenta llevada escrupulosamente porque la cantidad total que se invirtiera debía pagarse a medias entre los que decidieron que se hiciera: el Cabildo de la catedral de Valencia y la Cofradía de Ntra. Señora. La campana era para la torre llamada "el Micalet", construida a fines del siglo XIV.

Escondidos en las casi cuatrocientas partidas registradas se conservan detalles que permiten rehacer un poco la vida de aquel 1405: el salario de un oficial y de un peón, el precio de unos alimentos, la pobreza de medios, la carga máxima de un animal de tiro, los componentes de una comida, y, junto a ellos, los vocablos técnicos del material utilizado y de las operaciones que llevaron a la fundición de la campana. Las anotaciones, escuetas, no dan la razón del gasto, por lo que he tenido que descubrir para qué se destinó cada material y su papel en la operación.

La fundición de una campana hoy atraviesa las mismas fases que hace cientos de años; sólo se han introducido las ligeras modificaciones que han impuesto los cambios de la forma de vivir más que los avances científicos. El trabajo base de la fundición de una campana es la preparación del molde, que es doble: uno interno, llamado *macho*, y otro externo, llamado *capa*, que se superponen dejando entre los dos un espacio vacío, que ocupará el metal fundido, y que al solidificarse formará la campana. El macho por su cara externa fija el hueco o interior de la campana; y la capa, por su cara interna, la superficie exterior de la campana.

La preparación de estos moldes, que deben tener el mismo eje de simetría axial, obliga a montar un dispositivo de rotación que se mantiene fijo para toda la operación: el *torno*. A un eje de madera, clavado vertical en el suelo, se le ajusta en el extremo libre una pieza también de madera, llamada *terraja*, que puede girar libremente y en la que se van fijando, perpendiculares a la superficie del molde, unas láminas metálicas con un perfil que va cambiando a medida que se avanza en la tarea, perfil que corresponde a las líneas interior y exterior de las secciones de la campana y a la superficie externa de la capa.

La base común a todo el molde, y por lo tanto a cada una de sus partes, llamada *muela*, se forma con un cerco o plataforma de ladrillos, alrededor del eje, dejando entre ellos unos huecos que servirán de respiraderos para facilitar el tiro de las pequeñas hogueras que se encenderán en su interior para cocer los moldes, haciendo de chimenea el interior, vacío, del molde, saliendo el aire caliente por el hueco superior, en donde se coloca, al final, el molde de las asas.

El *molde*, que va a estar sometido a elevada temperatura cuando se cueza y cuando se realice la fundición, y sufrir unas operaciones mecánicas, debe tener una estructura sólida que impida la aparición de grietas y presentar cierta cohesión interna para que no se cuartee, desmenuce y desmorone. Estas cualidades se consiguen construyendo la médula del macho con fragmentos de ladrillos (tejas rotas) y arcilla, recubriéndolo todo con arcilla amasada con paja troceada, en la que se van insertando mechones de cáñamo y vueltas de alambre. En la parte interna se deja un hueco que permite trabajar la superficie y servirá, en su momento, de chimenea. La paja traba el cuerpo del macho. En el documento de 1405 se

anotan, seguramente para aumentar la trabazón, pelos de odre y borra, que se amasaban con la arcilla.

La paja, como tal, no aparece en el cuaderno de gastos; en su lugar figuran unas cargas de excrementos de bestias ("femtà de besties"); la paja cortada y quebrantada por la masticación, y medio digerida por los animales, es más flexible, flexibilidad que comunica a la masa del molde. En la actualidad la falta de animales de tiro en los lugares que están emplazadas las fundiciones ha hecho que los excrementos de los animales de tiro se sustituyan por paja, sin más.

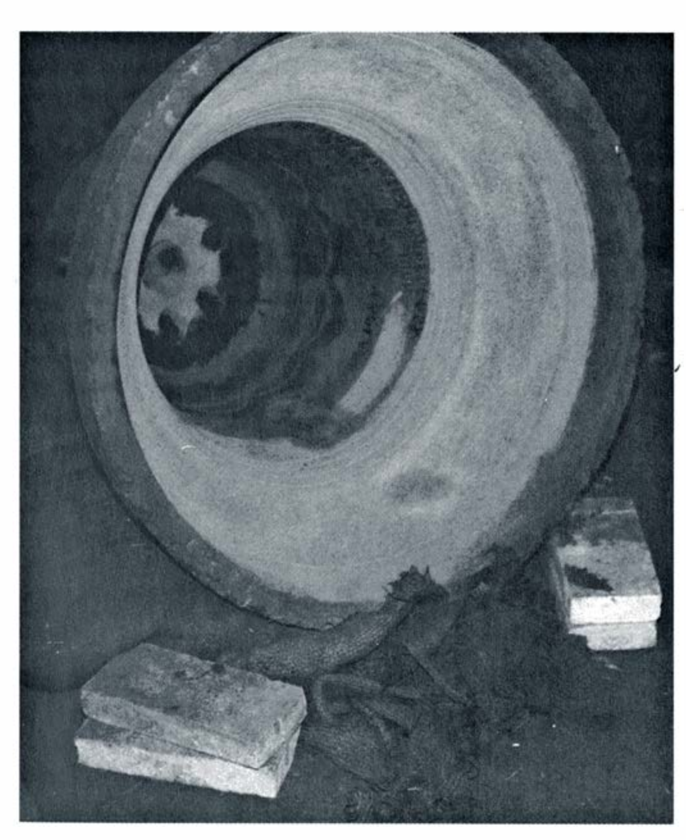
Sobre el macho, ya cocido, se coloca una capa formada por una lechada de ceniza, ceniza que previamente se cierne para separar las partículas gruesas y mal quemadas de la madera inicial, que se utilizó para obtenerla. La finalidad que se persigue con el encenizado es procurar que la parte interior de la campana, una vez fundida, quede lisa y no presente irregularidades debido a adherencias del macho. Químicamente tiene su explicación: el metal fundido, al ponerse en contacto con la parte "ácida" de la arcilla, principalmente dióxido de silicio, puede formar silicatos metálicos que quedarían fijados al resto del metal solidificado, apareciendo costras y adherencias. Si la parte del macho que va a estar en contacto con el metal fundido es fuertemente "básica", como lo es la ceniza, la posibilidad de formar silicatos adherentes con la masa del molde queda muy reducida.

Una vez seco y cocido el macho o molde interior, y seca la ceniza con la que se ha cubierto, se pasa a formar sobre él la *camisa* o *falsa campana*, llamada así porque es la parte que se sustituirá, cuando llegue la hora, por el metal fundido. La camisa se hace también de arcilla, a la que se agrega una buena parte de sebo, manteca u otra materia grasa. Aquí no se utiliza la paja para dar la cohesión. Al contrario de lo que ocurre con el macho, la falsa campana se eliminará en una operación posterior, por lo que debe poderse romper y desmenuzar con relativa facilidad. Esta es la principal razón de mezclar la arcilla con la grasa. Cuando se cueza el conjunto, la materia orgánica dispersa en la arcilla, al destruirse entre la arcilla, le dará una porosidad y soltura que hará que se quebrante fácilmente y pueda romperse.

A la falsa campana se le da forma con la plantilla correspondiente, distinta de la que se empleó para preparar el macho. Después de alisar su superficie, que será la superficie externa de la campana, se recubre con una suspensión de arcilla en agua a la que se añade cera. Sobre esta superficie encerada se fijan las letras de la inscripción, las decoraciones y los adornos, todo hecho, aparte, en cera, usando unos moldes de madera.

En 1405 se utilizaban tres clases de cera: cera engomada roja, cera blanca y cera nueva. La diferencia entre ellas residía en el precio –la cera engomada costaba el doble que la cera blanca, siendo la más barata la cera nueva– y en la elaboración, que a su vez dependía del destino. La cera engomada roja se destinó, según la anotación, para hacer la decoración ("imatges") y las letras; cabe presumir, pues, que el engomado fuera una resina que se le añadía para darle dureza. La cera blanca, purificada por fusión y filtración, serviría para mezclar con el sebo y cubrir la camisa; la cera nueva, posiblemente tal como se separa del panal, se usaba para fabricar las asas de la campana.

Fijadas las letras y la decoración, se aplican con pincel grueso, por toda la superficie de la camisa, unas manos de una papilla preparada con arcilla en suspensión en vinagre. Hasta hace poco se le añadía vino, claras de huevo y en alguna ocasión sangre de animal. Esta aplicación va formando una costra que recubre los relieves en cera de la falsa campana y una vez seca se pone sobre ella arcilla mezclada con paja y se



3. Negativo de la superficie en la parte interior de la capa



4. Colocación de la capa sobre el macho dejando el hueco para el metal

forma la masa principal del molde exterior, con grueso suficiente; desaparece la línea característica de la campana y toma el todo una forma troncocónica.

La forma exterior del molde se le da con la pieza correspondiente de la plantilla, que se hace girar, como en las operaciones anteriores, alrededor del eje de la campana. También este molde se apoya en la amplia base preparada al principio, la muela, alrededor del eje del torno.

La presencia de las claras de huevo y vino o del vinagre actual en la papilla, materia y compuestos orgánicos, aseguraba que, una vez seco y cocido el molde, y antes de fundir, se desprendiera con facilidad esta parte del molde de la superficie exterior de la camisa, sobre el que está puesto. En la masa del molde, que lleva sus refuerzos de alambre y fibras de cáñamo, se insertan unas piezas de hierro, aros, para evitar que el conjunto se abra, y unos garfios que servirán, en su momento, para levantar la capa. A la arcilla de la capa se le mezclaba de nuevo una sustancia grasienta. Este molde exterior se seca y se cuece: la cera de los relieves colocados en la superficie exterior de la camisa o falsa campana se funde y es absorbida por la arcilla que estaba en contacto con ella.

Aparte se preparan las asas de la campana, haciendo un negativo en yeso o madera, rellenándolo con cera y recubriendo el positivo obtenido en cera con arcilla amasada con paja y reforzada con cáñamo. Este molde está abierto por su parte inferior para que pueda formar un todo con el molde del cuerpo superior de la campana; por la parte de arriba se dejan unos canales a través de los cuales entrará el metal fundido y saldrá el aire desplazado en la colada. El molde de las asas se cuece en un horno y se recoge la cera fundida; queda el molde vacío y preparado para colocarlo y unirlo a la parte superior del molde de la campana. El conjunto se cubre con una capa de arcilla, se cuece y se coloca el molde entero en un hoyo abierto al pie del horno, cerca y por debajo de la piquera u orificio por donde saldrá el metal fundido.

Una vez bien asentado el molde en el hoyo, se procede a separar del macho la capa, para lo cual se ejerce, por medio de un sistema mecánico de elevación, una tracción sobre los garfios insertos en la capa, separación que viene facilitada por el tratamiento de huevo, vino, etc., que se dio a las superficies que iban a estar en contacto. Sobre el molde del macho aparece la camisa, que a continuación es troceada y separada del mismo. La parte interior de la capa, que presenta el negativo de la superficie exterior de la futura campana, se repasa por el maestro campanero por si alguna letra o motivo decorativo ha sufrido desperfecto.

Limpias las superficies, exterior del macho e interior de la capa, se recubren de una lechada de ceniza (como se hizo en una operación anterior para que el metal al solidificar no se adhiera a la arcilla) y se coloca de nuevo la capa sobre el macho, quedando entre los dos moldes el hueco que llenará la fundición. Terminada esta operación de preparación del molde, se rellena el hoyo con la tierra extraída, y se apisona a su alrededor.

En la Edad Media no hubo, por lo general, instalaciones industriales fijas, salvo las que exigían disponer de una fuente de energía no transportable (molinos o fraguas). Los fundidores de campanas, y los impresores, improvisaban su taller allí donde se reclamaba su servicio. Operación paralela a la preparación del molde para la campana la constituía, pues, la construcción de un horno para fundir el metal. De acuerdo con nuestro documento, se procedió a construir un horno sirviéndose de materiales corrientes, aunque para la parte interior se utilizó una tierra blan-

ca ("terra de ferrers"), posiblemente dolomita.

De los pesos de cobre y estaño que aparecen comprados se deduce la composición de la fundición: 76,6 por ciento de cobre y 23,3 por ciento de estaño. Los metales, troceados, se colocaron en el horno que se calentó con madera de pino y olivo, y carbón, activando la combustión con unos fuelles, movidos ininterrumpidamente por operarios, que se relevaban en el trabajo. La masa fundida se removía con unas varas largas de madera, y la escoria que se forma y flota se extraía con unos cucharones, de madera de pino, carbonizada superficialmente.

Los fundidores tenían especial interés en quedarse con las escorias. En muchos contratos de fundiciones de la época se especifica que las escorias quedarán en poder del fundidor, y se negociaba como si fuera algo de especial valor. Puesto que la escoria, formada por silicatos e impurezas, retiene poco metal incluido, hay que pensar que el interés por la misma debía ser otro que el metal que pudiera recuperarse. He llegado al descubrimiento de dicho interés. Entre los fundidores está muy extendida la creencia de que la sonoridad de una campana se mejora añadiendo a la masa fundida plata. Sin embargo, cuando se analizan los bronces de campanas antiguas, la proporción de plata que aparece es insignificante. Sí suele hallarse en la composición algo de plomo; dada la solubilidad de la plata en el plomo fundido, y la fácil formación de escoria, la plata puede acumularse en ella.

La plata añadida a la masa fundida, e introducida en el crisol por devoción religiosa y manifiesto desprendimiento, la concentraba en la escoria el plomo o quizás el mismo estaño; de esta manera se recuperaba después en una operación privada por el maestro fundidor, que agregaba así una nueva ganancia a su trabajo. (En la fundición se admitía una merma, en torno al ocho por ciento en peso, de metal inicial.)

Realizada la fundición, la experiencia había enseñado que un enfriamiento rápido podía provocar defectos en la pieza. La campana, que debió pesar algo más de dos mil seiscientos kilos, tardó en fabricarse unos cuarenta días, desde el sábado 21 de febrero de 1405 al jueves 2 de abril, en que quedó fuera del molde y se limpió su superficie con cepillos y vinagre. Trabajaron en ella el maestro campanero Guillem Martí de Paris y su ayudante, tres oficiales y cuatro peones. Los peones cobraron dos sueldos y medio y los oficiales cuatro sueldos y medio, diarios. El maestro campanero cobró por su trabajo 935 sueldos, es decir, 85 florines.

En el contrato de fundición se especificaba que la campana debería quedar en perfectas condiciones y que si por cualquier causa resultaba defectuosa el maestro campanero debería volverla a fundir a su costa. La bondad de la fundición, y por tanto la aceptación de la campana, la fijaban unos expertos que la sometían a unas pruebas de sonoridad. La prueba la llamaban la prueba del hilo y consistía en ceñir ajustado un hilo a la campana y golpearla repetidamente. Si el hilo se rompía era señal de que la campana estaba agrietada.

Esta prueba tan sencilla tiene su explicación física. En la vibración de una campana, compleja en sí, se forman ondas estacionarias. Unas de ellas se forman en planos perpendiculares al eje de la campana, lo que hace que la superficie de ésta se deforme por la vibración, pero sin que varíe la longitud del desarrollo en el plano, es decir, varía la forma pero no el perímetro. Si tiene una grieta, al estar libre el lugar donde se ubica, origina un vientre en la vibración, lo que llevará consigo la variación de la línea envolvente que forma el hilo, y se producirá su rotura. (José Sánchez Real.)

Base molecular del cáncer

Los cánceres humanos comienzan por la acción de oncogenes, versiones alteradas de genes normales. En un caso, la alteración es una mutación puntual, que cambia un solo aminoácido de la proteína codificada

Robert A. Weinberg

Un tumor maligno es un gran conglomerado de células cancerosas, descendientes todas ellas de una célula "fundadora". Tal célula ancestral era antes normal, y desempeñaba su función regular en algún tejido. Pero sufrió un cambio crítico y, como resultado del mismo, empezó a dividirse y proliferar en respuesta a alguna orden interna, haciendo caso omiso a los estímulos externos que normalmente regulan el crecimiento celular. Esta única célula generará los miles de millones de células, igualmente alteradas, que constituyen la masa tumoral.

Para saber cuál es el factor determinante del desarrollo de un cáncer debemos conocer qué alteraciones ha sufrido la célula fundadora. ¿Por qué las células ignoran los límites impuestos en el crecimiento normal? En los últimos años han empezado a esbozarse algunas respuestas. En los cromosomas de las células tumorales se han descubierto genes implicados en el desarrollo canceroso. Estos genes, a menudo llamados oncogenes, constituyen la fuerza motriz determinante del crecimiento incontrolado de muchas células cancerosas. Dichos genes se activan durante la conversión de la célula fundadora normal en célula cancerosa. Una vez activados, conducen a las células hacia ese comportamiento anormal conocido como estado canceroso.

Pleiotropía

El comportamiento anormal de las células cancerosas se caracteriza por la presencia de muchos rasgos específicos. El crecimiento incontrolado es el más evidente. Además, las células cancerosas se suelen distinguir de las normales por su forma. No respetan las reglas territoriales que confinan a las células normales en tejidos concretos. Muchas absorben cantidades anormalmente altas de azúcares. Poseen una tasa elevada de metabolismo anaeróbico (proceso de obtención de energía independiente del oxígeno). La membrana externa de las células cancerosas difiere de la que tiene una célula normal, y posee además antígenos tumorales que les confieren propiedades inmunológicas específicas.

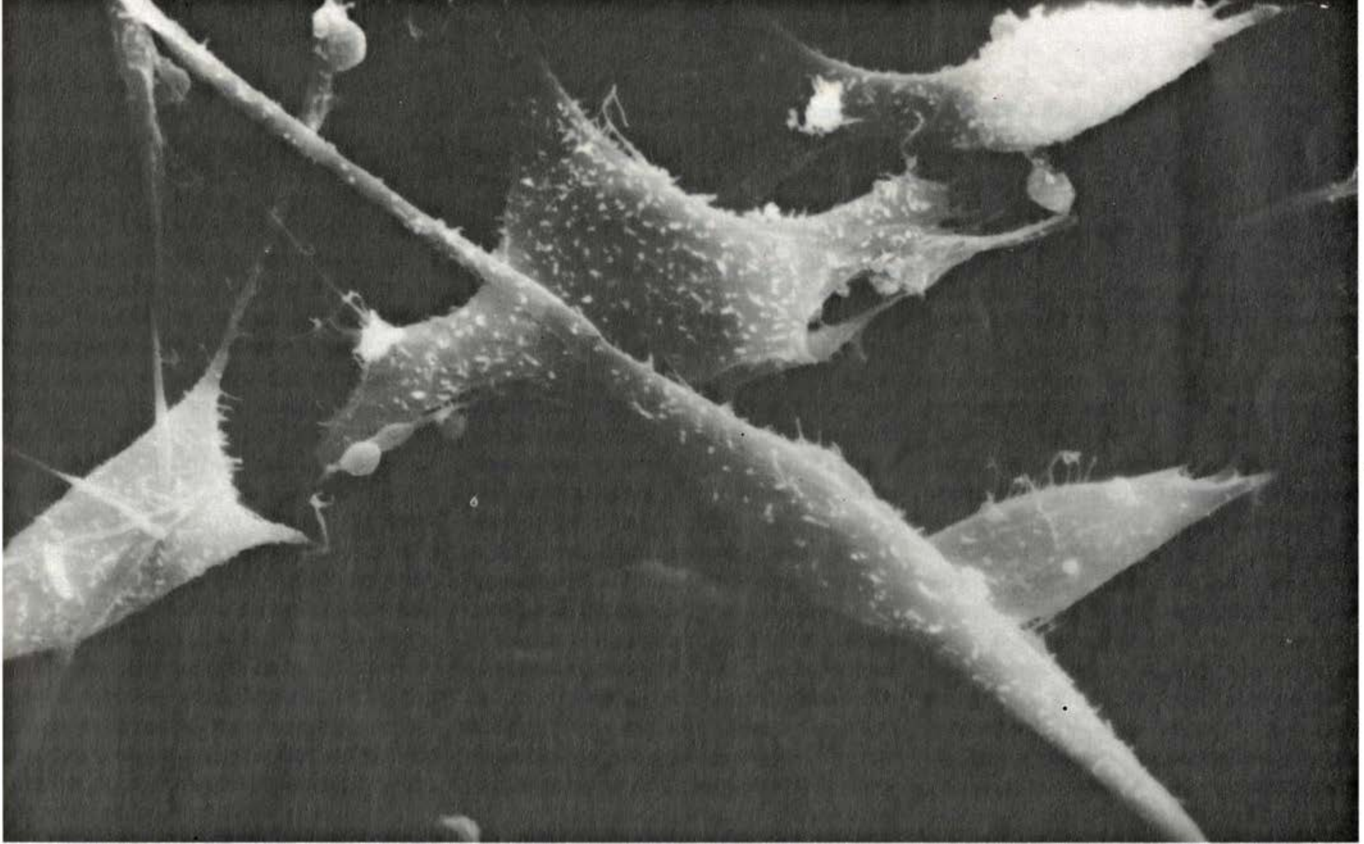
Esto es sólo el principio de una larga lista, cuya longitud, sumada a la propia complejidad del fenotipo canceroso (conjunto de características estructurales y funcionales que definen una célula cancerosa), plantea interrogantes importantes. ¿Qué características de entre ellas son esenciales para alcanzar el estado canceroso, y cuáles secundarias? Si una célula cancerosa posee cien rasgos específicos, ¿se adquieren todos ellos uno a uno, en sucesivas etapas discretas, durante la carcinogénesis? ¿Pasa la célula por cien estadios en el curso de su evolución desde el estado normal, implicando cada estadio un cambio diferente? ¿Hay, por el contrario, un mecanismo "pleiotrópico" más simple? Pudiera ser que un único elemento celular se activara y posteriormente determinara simultáneamente un gran número de cambios en el fenotipo.

Afortunadamente parece que priva la simplicidad. Los primeros indicios que apuntaban en este sentido se remontan a los estudios realizados hace unos 20 años con unos pequeños virus de ADN que inducen cáncer en algunos animales. Marguerite Vogt y Renato Dulbecco, que a la sazón trabajaban en el Instituto de Tecnología de California, lograron la primera transformación de células normales en cancerosas dentro del ámbito de un laboratorio, en una placa con un cultivo de tejidos. Incubaron el virus polioma en fibroblastos (células de tejido conectivo) de embriones de hámster, mantenidas en placas de Petri; observaron que en los cultivos infectados aparecían colonias de células alteradas que se amontonaban unas sobre otras, en vez de formar una capa única, o monocapa, como ocurría con los fibroblastos normales. Cuando las células alteradas se inocularon en ratas jóvenes provocaron la aparición de tumores. En otras palabras, manipulando células normales en una placa de Petri se produjeron células tumorales; se había logrado inducir la "transformación" de células normales en tumorales. La inducción del cáncer dejaba de ser un proceso misterioso que sólo tenía lugar en los inaccesibles tejidos de un animal.

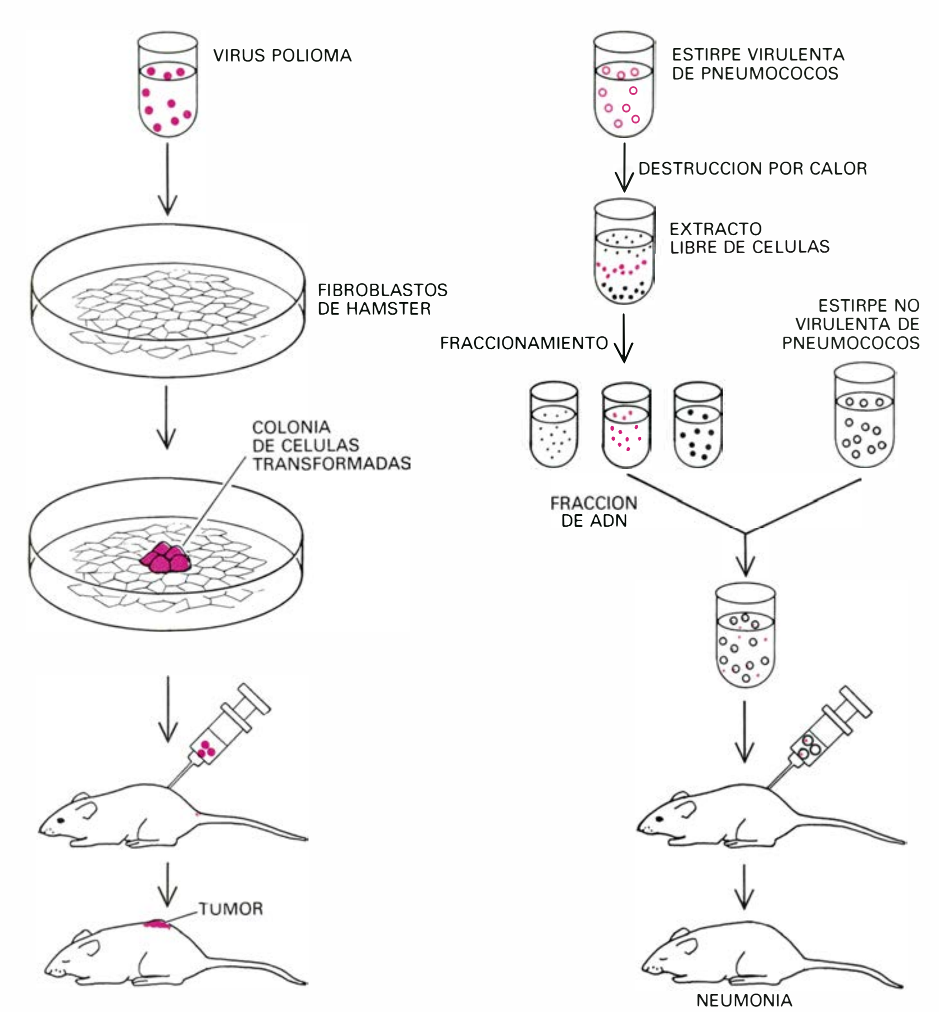
Las células transformadas por virus presentaban muchas de las características asociadas con el fenotipo canceroso. Dulbecco y otros demostraron que la transformación la inducían genes víricos, segmentos de información genética que el virus introduce en la célula durante el proceso de infección. Esta información está contenida en las moléculas de ADN de las partículas víricas. Cuando un virus cancerígeno infecta una célula animal, introduce una cantidad muy pequeña de ADN, quizás una millonésima parte de la que contienen los propios cromosomas de la célula. Unos pocos genes víricos habían logrado, por tanto, inducir docenas de cambios en la estructura y comportamiento general de la célula. Resultaba obvio que se producía un efecto pleiotrópico en la transformación.

Base genética

Este trabajo pionero condujo a otro descubrimiento relacionado con la base molecular de la transformación. Si durante la proliferación de las células cancerosas se perdían o se inactivaban experimentalmente los genes víricos, las células revertían a un estado normal. No sólo se requerían genes víricos para iniciar la transformación, sino que era necesaria su presencia y continua acti-



1. TRANSFORMACION DE CELULAS NORMALES por un oncogén, según se observa en la micrografía de barrido realizada por Erika A. Hartwieg y Jonathan A. King, del Instituto de Tecnología de Massachusetts. Se trataron fibroblastos normales de ratón (*arriba*) con ADN extraído de la línea de células humanas cancerosas denominada *EJ*, derivada de un carcinoma de vejiga. El ADN del tumor "transfectaba" a las células de ratón (esto es, las infectaba y las transformaba), dando lugar a colonias de células cancerosas (*abajo*). Mientras que las células normales son planas y guardan cierta distancia entre sí, las transformadas se incurvan, proliferan y se amontonan unas sobre otras. Las células se fijaron con glutaraldehído y osmio, se deshidrataron, se secaron cuidadosamente y se cubrieron con una delgada película de oro y paladio. El aumento es, en ambas micrografías, de 2000 diámetros.



2. TRANSFORMACION CELULAR CON ADN VIRICO, según demuestra un experimento (*izquierda*) realizado por Marguerite Vogt y Renato Dulbecco. Trataron fibroblastos de hámster con virus polioma. El pequeño genoma (dotación genética) vírico se basta, una vez introducido en una célula animal, para dar lugar a colonias de células transformadas, que presentan muchas características alteradas. La inyección de células transformadas en ratón origina tumores. No obstante, se necesitaba una demostración directa del papel desempeñado por el ADN celular en la transformación. Para hallarla se utilizó una aproximación experimental basada en el histórico experimento de Oswald T. Avery y sus colegas, que, en 1944, confirmaron que el ADN era el material genético (*derecha*). Tras fraccionar un extracto libre de células de pneumococos virulentos, demostraron que sólo la fracción correspondiente al ADN era capaz de volver virulenta una estirpe que hasta entonces no lo era. (Los dibujos son obra de Bunji Tagawa.)

vidad para mantener el fenotipo canceroso.

Así las cosas, era lógico especular con la posibilidad de que en otras células cancerosas transformadas por otros agentes no víricos operasen mecanismos análogos. ¿Llevarían todas las células cancerosas un pequeño número de genes cuya actividad fuese necesaria para dirigir el complejo comportamiento de las mismas? Ello implicaría obviamente una base genética para el cáncer: que el estado canceroso dependiese de ciertos genes y no de otros sistemas reguladores "epigenéticos" que afectasen a la célula.

Veinte años atrás, esta idea estaba lejos de ser aceptada sin reservas, aun cuando desde hacía bastante tiempo existían otros indicios de que ciertos genes podían hallarse implicados en el cáncer. Otros investigadores habían acumulado datos relativos a carcinógenos no víricos, entre los que se incluían varias formas de radiación y una amplia gama de agentes químicos inductores de transformaciones celulares. Tales carcinógenos no aportaban información genética nueva a la célula, sino que alteraban la ya existente, es decir, a los genes de las células afectadas. Muchos carcinógenos actúan dañando al ADN y provocando mutaciones génicas. De nuevo, los genes parecían desempeñar un papel central en la carcinogénesis.

"Parecían desempeñar" es todo lo que podía afirmarse, ya que no había pruebas directas. Hacía falta un experimento que demostrase directamente la presencia en la célula cancerosa de genes alterados, responsables de su comportamiento aberrante. La forma de demostrarlo era analizar el complejísimo ADN celular y no el ADN relativamente simple de un virus cancerígeno. Una propuesta en este sentido se basaba en el histórico experimento que, hace 40 años, demostró por primera vez el papel central del ADN como portador de la información genética. Oswald T. Avery, Colin M. MacLeod y Maclyn McCarty, del Instituto Rockefeller de Investigaciones Médicas, estudiaban dos estirpes de la bacteria pneumococo, una de las cuales provocaba una neumonía letal en ratones. ¿Cuál era la base molecular que justificaba la diferencia entre ambas estirpes? Los investigadores analizaron un extracto de la bacteria virulenta y purificaron un "principio transformante", capaz de convertir las bacterias no virulentas en organismos activamente patógenos. Concluyeron que el principio transformante era el ADN. Cuando se transferían las moléculas de ADN se transmitía, de una célula bacteriana a otra, la información genética para la virulencia. Se demostraba así que el ADN portaba la información genética para un carácter discreto.

Transferencia de genes

Hace cinco años, Chiaho Shih realizó, en mi laboratorio del Instituto de Tecnología de Massachusetts, un experimento similar de transferencia de genes. Pretendía resolver el siguiente interrogante: ¿podía transmitirse el carácter canceroso desde una célula de mamífero hasta otra, transfiriendo solamente moléculas de ADN? Unos años antes, investigadores holandeses habían puesto a punto una técnica que permitía transferir genes entre células de mamífero. En dicha técnica se utilizan cristales de fosfato cálcico para retener moléculas de ADN. Tales cristales facilitan la entrada del ADN en las células receptoras; una vez dentro de ellas, el ADN transferido se integra en el ADN cromosómico.

Shih extrajo ADN de células cancerosas, concretamente de una línea celular de fibroblastos de ratón, transformados tras una exposición al metilcolantreno, un carcinógeno químico. El ADN de la célula donante se coprecipitó con fosfato cálcico, y se añadió luego a cultivos de fibroblastos de ratón no transformados, células éstas que pertenecían a una línea de laboratorio denominada *NIH*3*T*3. A las dos semanas de añadir el ADN donante se detectaron colonias de células transformadas en

los cultivos de *NIH*3*T*3. Las colonias transformadas, vistas al microscopio, eran muy parecidas a las inducidas mediante la infección con virus. Estas células transformadas, inoculadas en ratones, fueron, a su vez, capaces de producir tumores.

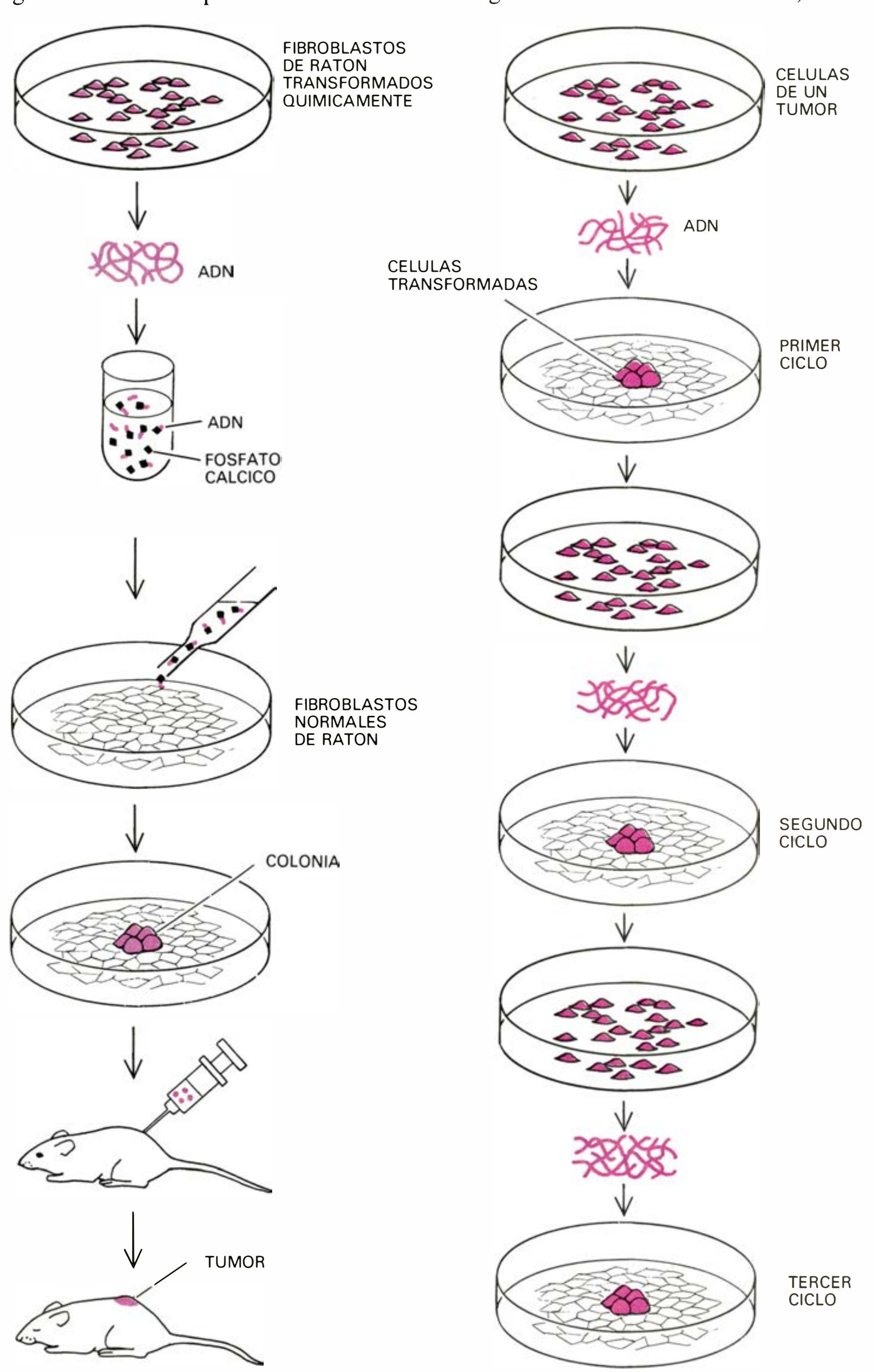
No había duda: la molécula de ADN había transferido la información para el cáncer de una célula a otra. La expresión de parte, al menos, del fenotipo canceroso se debería directamente a la información presente en el ADN de las células de mamífero que habían sido transformadas por un carcinógeno químico. Cuando se transfería ADN de células normales no se observaba ningún efecto. Ello demostraba que algunas secuencias del ADN de las células cancerosas donantes diferían de sus análogas presentes en el ADN de las células normales. Sospechábamos que tales diferencias quizá se debieran a mutaciones en el ADN de las células cancerosas donantes, causadas por el metilcolantreno.

Podía cuestionarse la extrapolación de las primeras observaciones, puesto que el ADN donante procedía de fibroblastos de ratón transformados químicamente, que a su vez habían sido introducidos en células no transformadas del mismo tipo. Cabía la posibilidad de que la transformación mediada por ADN que estábamos observando fuese característica exclusiva de las células donantes y receptoras de este tipo especial, y que se obtuvieran resultados diferentes con otros tipos celulares. Se demostró que este no era el caso. Investigadores de nuestro y de otros laboratorios repitieron el experimento inicial, con ADN procedente de distintos tipos de células cancerosas y encontraron la misma actividad transformante. Entre los tumores humanos utilizados como fuente de ADN donante había carcinomas de vejiga, colon y pulmón, así como fibrosarcomas, neuroblastomas e incluso leucemias.

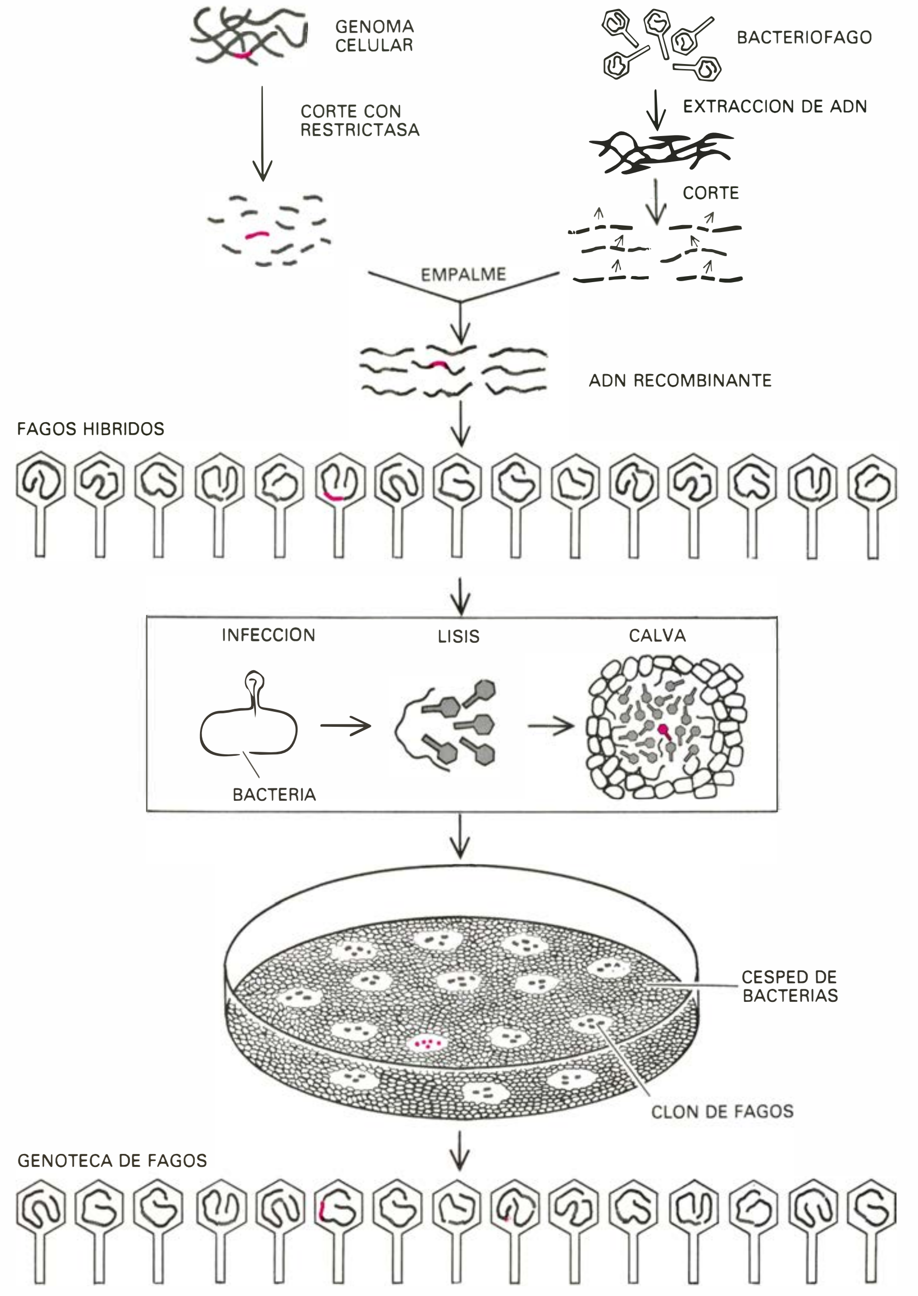
Debía haber, pues, algo en común entre los fibroblastos de ratón transformados químicamente y las células de todos aquellos tumores humanos espontáneos, ya que el ADN de todas ellas tenía el mismo efecto cuando se transfería. Más aún, si el ADN extraído de un carcinoma de colon podía transformar fibroblastos de ratón, no había duda de que las secuencias transformantes del ADN del tumor funcionarían en células de distinta especie y tejido. Por último, el ADN transformante introducido en la célula mediante la técnica de transferencia parecía actuar pleiotrópicamente, ya que las células transformadas por dicho ADN presentaban ahora varias de las características típicas de las células cancerosas. Como en el caso del ADN del virus polioma, una cantidad limitada de la información genética donante provocaba alteraciones en el comportamiento de la célula receptora.

El principio transformante

Los experimentos de transferencia de genes nos revelaban mucho, mas



3. POR TRANSFERENCIA GENICA (*izquierda*) se demostró que el ADN de células cancerosas determinaba características cancerosas. Se extrajo ADN de células de ratón, previamente transformadas mediante un carcinógeno químico. Se coprecipitó dicho ADN con fosfato cálcico, que facilita su entrada en células normales de ratón. Las células normales transfectadas originaban colonias de células transformadas. La inyección de las células transformadas en ratones daba lugar a un tumor. Con experimentos de transferencia (*derecha*) se vio que el ADN extraído de células de un tumor humano transformaba células normales de ratón. Las células transformadas se cultivaron a gran escala, se extrajo su ADN y se volvió a utilizar éste para transformar células de ratón. El proceso se repitió por tercera y cuarta vez. Sólo segmentos cortos de ADN podrían sobrevivir a las diferentes extracciones y coprecipitaciones. La existencia de transformación demostraba que el agente transformante debía residir en uno de esos segmentos.



4. GENOTECA obtenida clonando genes en el bacteriófago lambda, un virus. Tanto el genoma celular como el del fago se cortan con una restrictasa. Se elimina un segmento del ADN del fago, y el resto se une a un fragmento de ADN celular mediante el enzima ligasa de ADN. El ADN recombinante se empaqueta en la cubierta proteica del fago lambda. Cuando un fago híbrido infecta la bacteria *Escherichia coli*, el fago se multiplica y lisa la célula (la hace estallar); los fagos descendientes salen e infectan y matan a las células de alrededor, formando un halo o "calva". Cuando en una placa de cultivo se inocula un "césped" bacteriano con gran número de fagos híbridos diferentes, cada halo presente en el césped está formado por un solo clon de fagos, descendientes todos ellos del fago que infectó la bacteria. Cada clon lleva un fragmento de ADN celular diferente. Se plantea entonces el problema de identificar el clon portador del gen buscado (*color*). Normalmente se logra sondeando los clones con ADN o ARN de secuencia parecida al gen deseado; mas para la búsqueda del ADN transformante no se contaba con sonda apropiada.

para ahondar en ese terreno era necesario aislar y caracterizar la secuencia o secuencias específicas responsables de la transformación, ya que ni siquiera conocíamos el aspecto, quizás más elemental, del "principio transformante" del ADN del tumor; dicho de otra manera, ¿constaba tal principio de un solo segmento de ADN o de varios elementos genéticos independientes cuya cooperación se exigiría para transformar una célula? Si se debiese a un solo segmento, ello sugeriría la existencia, como responsable, de un solo gen, similar a los muchos que hay en el genoma celular (dotación genética completa).

La posibilidad de que el principio transformante se localizase en un solo segmento de ADN contaba ya con un apoyo experimental. El ADN extraído de células cancerosas donantes y añadido posteriormente a fibroblastos de ratón inducía, lo sabíamos, la aparición de colonias de fibroblastos transformados. Se aislaron las células de una de las colonias transformadas y se cultivaron a gran escala. Cuando se extrajo el ADN de este cultivo y se añadió de nuevo a células normales, volvían a aparecer colonias de células transformadas. El proceso se podía repetir hasta un tercer o cuarto ciclo de transferencia. Cabía deducir de ello que el principio transformante, cualquiera que fuese su naturaleza, residía en un único segmento de ADN, que era capaz de resistir las sucesivas transferencias. Las diferentes extracciones y consiguientes manipulaciones hacían que las largas cadenas del ADN celular se rompiesen en miles de pequeños segmentos, de los cuales sólo unos pocos probablemente entrasen en una célula concreta de ratón. Si la actividad transformante dependiese de la cooperación de múltiples e independientes elementos, seguramente se hubiese perdido durante las transferencias, ya que tales elementos se habrían disociado unos de otros.

El problema, por tanto, consistía en hallar el segmento transformante y estudiarlo. La propia estructura del ADN determinaba, en cierta forma, el diseño de los experimentos a realizar. La molécula de ADN es una doble hélice formada por dos cadenas de nucleótidos que se emparejan, y que se caracterizan cada uno de ellos por la presencia de uno de los cuatro grupos químicos llamados bases: adenina (*A*), guanina (*G*), timina (*T*) y citosina (*C*). El mensaje genético está contenido en una secuencia formada por combinaciones de los cuatro nucleótidos de ADN. La expresión de dicho mensaje tiene lugar cuando una de las cadenas se transcribe en otro ácido nucleico, el ARN, y éste se traduce en una cadena de aminoácidos, esto es, una proteína. Sospechábamos que el carcinógeno metilcolantreno había inducido mutaciones en el ADN de la célula cancerosa donante y que en eso se diferenciaba de la normal.

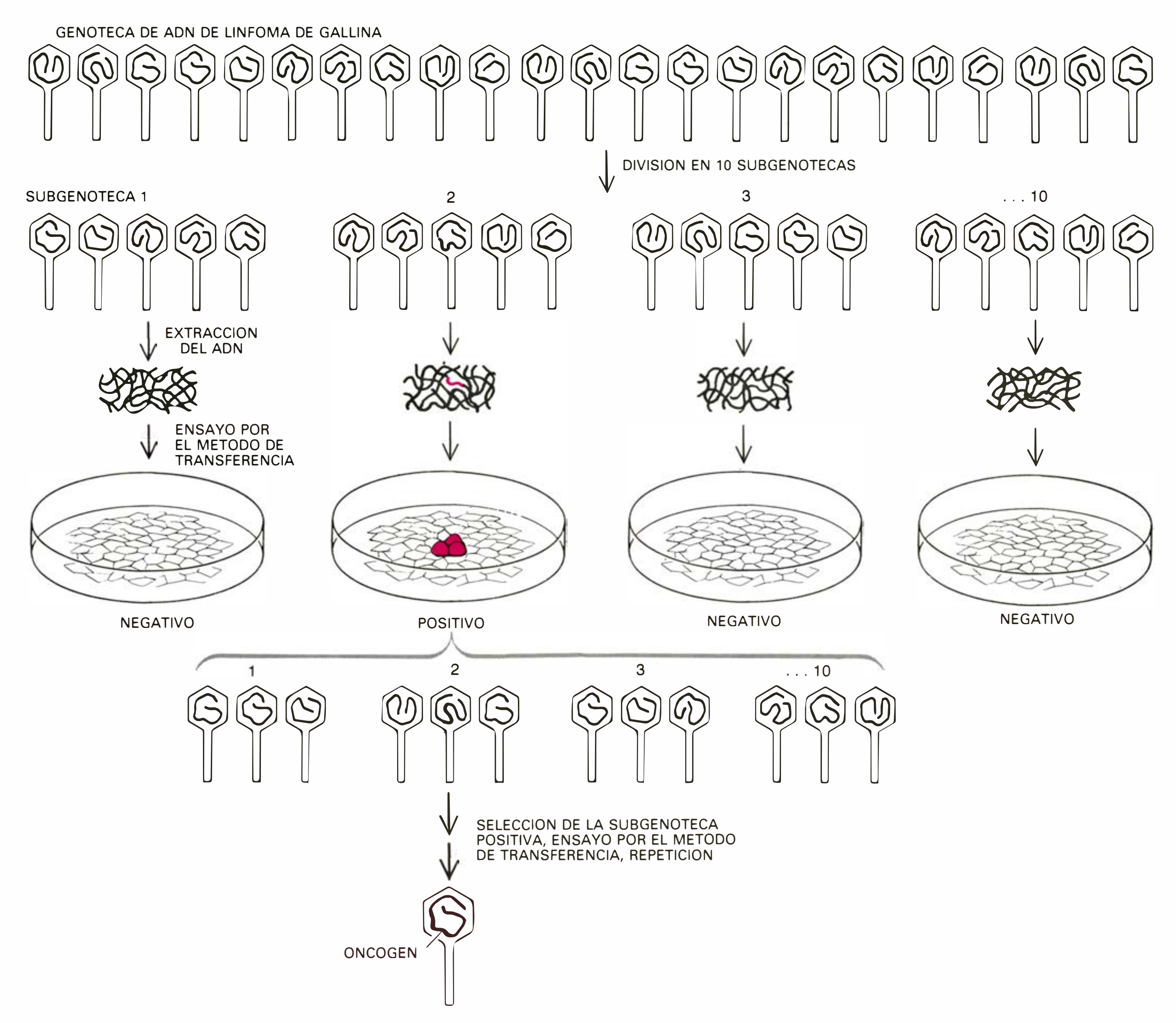
El genoma celular consta de unos 6000 millones de pares de bases de ADN. Por término medio, un gen contiene de 5000 a 10.000 pares de bases. El reto no era otro que aislar un segmento genético, portador del principio transformante, entre una cantidad un millón de veces mayor de material genético que no nos interesaba para nada.

Las nuevas técnicas de clonación de genes lo plasmaron en la realidad. Con dichas técnicas podemos aislar un segmento genético específico del genoma celular y multiplicarlo (amplificarlo) para conseguir muchos miles de copias idénticas. Así se estudia un gen puro, libre de la contaminación que supone la enorme y compleja maraña formada por su "ambiente genético" normal.

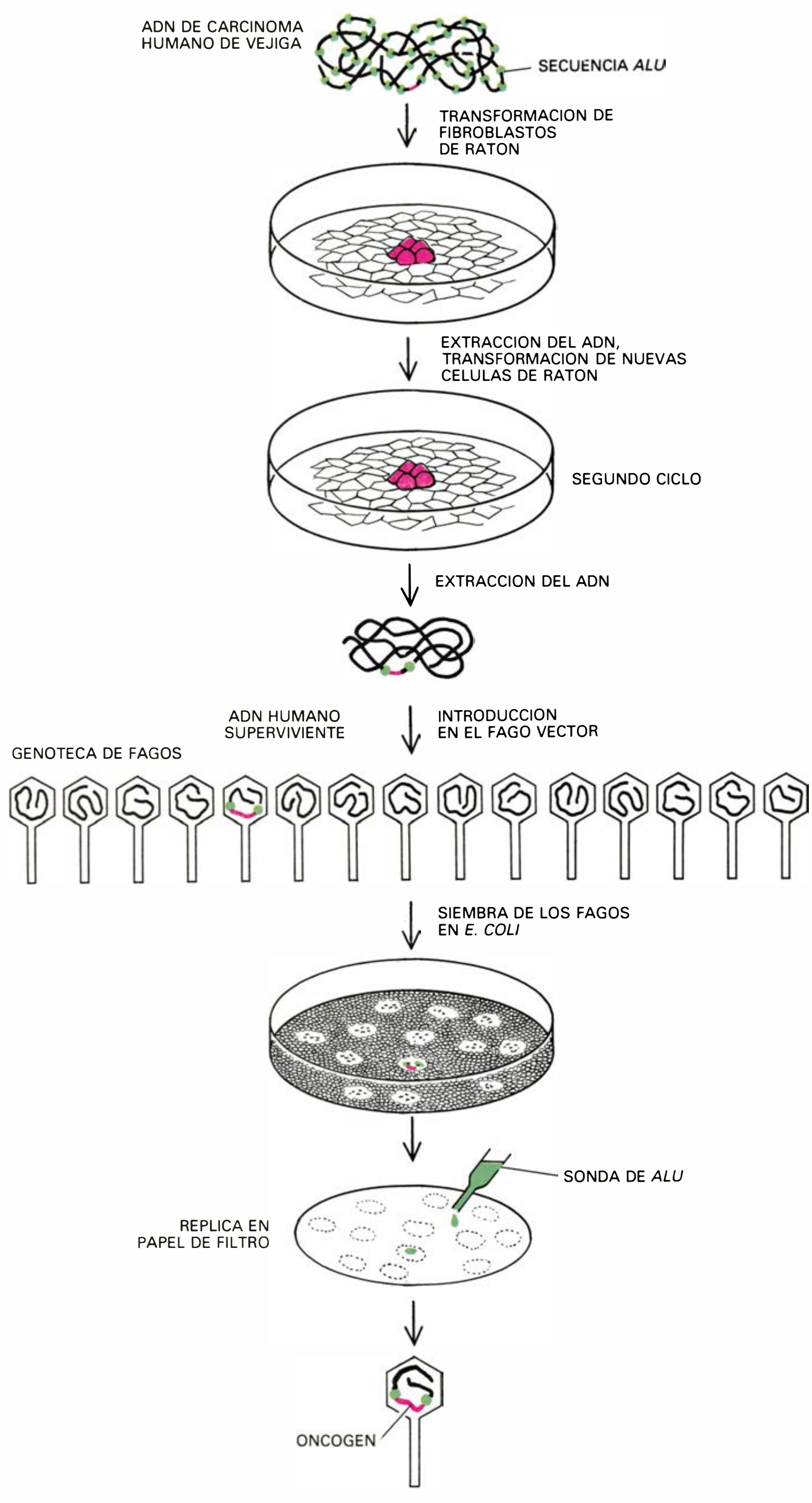
Las diferentes estrategias de clonación de genes, tanto las que utilizan plásmidos bacterianos (pequeños círculos de ADN extracromosómico) como las que se sirven de bacteriófagos (virus que infectan a bacterias), tienen una lógica común. El genoma celular se fragmenta en varios cientos de miles de segmentos. Cada uno de estos segmentos se inserta en un vector, o "vehículo", que en ocasiones es el pequeño genoma de ADN del bacteriófago lambda, un bacteriófago que infecta la bacteria *Escherichia coli*. Cada fago híbrido resultante, además de algunos de sus propios genes, lleva inserto un segmento de ADN "intruso". La colección formada por los miles de fagos híbridos recibe el nombre de genoteca, ya que en su conjunto encierra toda la información genética del genoma celular: todos los segmentos del genoma original se hallan presentes en uno o más de los fagos que componen la genoteca.

Estos fagos pueden multiplicarse (amplificarse) por separado, cultivándolos en *E. coli*. Cuando se inoculan los fagos sobre un "césped" de *E. coli*, en una placa de cultivo, cada uno de ellos infecta una célula bacteriana y fabrica muchas copias de sí mismo. Los descendientes del fago matan la célula, se liberan e infectan células vecinas, produciéndose, en pocas horas, miles de fagos descendientes que originan un halo o "calva" en el césped bacteriano. La población de fagos contenida en cada uno de los halos es clonal: descienden todos de un solo antecesor. A partir de un clon de fagos se aíslan miles de copias del ADN insertado, con lo que se obtiene un clon de dicho segmento de ADN.

El principal problema a resolver en



5. ESTRATEGIA DE CERCO utilizada por el grupo de Geoffrey Cooper para aislar el ADN transformante. Construyeron una genoteca, formada por unos 200.000 clones de fagos, a partir de ADN de células de linfoma de gallina. Dividieron la genoteca en 10 subgenotecas y ensayaron el ADN de cada una de ellas, utilizando la técnica de transferencia génica. Si la transferencia daba resultado positivo, la subgenoteca correspondiente, que necesariamente debía incluir el ADN buscado, se volvía a dividir en otras 10 subsubgenotecas, que se sometían de nuevo al ensayo. Se estudiaban sucesivamente los subgrupos que daban positivo, hasta aislar un clon de fagos que daba actividad transformante. A partir de él se aisló el oncogén de linfoma.



6. SONDA PARA ADN HUMANO; su utilización permitió a Chiaho Shih encontrar el gen transformante de un cáncer de vejiga humano. Desperdigadas por el genoma humano existen breves secuencias de ADN llamadas *Alu* (*verde*). Se transformaron las células del ratón con el ADN del tumor; el ADN de las células transformadas se utilizó para transformar nuevas células de ratón en un segundo ciclo. El ADN de las células sucesivamente transformadas incluiría sólo una pequeña cantidad de ADN humano, aquélla directamente asociada con la transformación (*rojo*). Con el ADN de las últimas células transformadas se construyó una genoteca, que se ensayó en *E. coli*. El ADN de los halos resultantes se transfirió a papel de filtro, y a éste se añadió una sonda con secuencias *Alu* marcadas con un isótopo radiactivo. La autorradiografía revelaba la localización del halo que contenía el ADN que hibridaba con la sonda. Se comprobó la capacidad transformante del clon de fagos del halo y se demostró la presencia del oncogén en su ADN.

los experimentos de clonación es el de identificar, partiendo de una genoteca, los contados clones de fagos que llevan inserto el ADN que nos interesa. Para hacerlo, lo mejor es utilizar una "sonda", que es un segmento de ADN monocatenario de secuencia nucleotídica similar, en su totalidad o en parte, a la del gen que se quiere clonar. Como las dos cadenas de una molécula de ADN son complementarias (*A* empareja con *T* y *G* con *C*), la sonda forma un "híbrido" específico con las cadenas del ADN buscado. Añadiendo una sonda marcada con un isótopo radiactivo a los halos de fagos, se localizará con precisión cuál de ellos contiene el ADN que buscamos. La detección de dicho halo se consigue mediante una técnica denominada autorradiografía.

Pero no disponíamos de sondas específicas para identificar las secuencias transformantes. Lo único que conocíamos del ADN que buscábamos era su actividad biológica, es decir, la capacidad de inducir transformación. Se necesitaban, por tanto, nuevas técnicas.

Escudriñando una genoteca

Tres laboratorios emprendieron la clonación, utilizando estrategias diferentes. Uno de los grupos, el dirigido por Geoffrey M. Cooper, del Instituto del Cáncer Dana-Farber, asociado con la Facultad de Medicina de Harvard, prescindió de sonda. Construyeron una genoteca utilizando ADN de células de linfoma de gallina. Por experimentos de transferencia génica se sabía que dicho ADN llevaba una secuencia transformante. La genoteca se subdividió en diez partes, y se purificó el ADN de los fagos que componían cada una de las diez subgenotecas. Mediante transferencias ensayaron la capacidad de dicho ADN para transformar fibroblastos de ratón. Las subgenotecas que daban resultado positivo se volvían a subdividir en otras diez, cada una de las cuales se examinaba nuevamente siguiendo la misma metodología. Esta búsqueda rutinaria basada en una estrategia de cerco exponencial llevó hasta un único clon de fagos con una fuerte actividad transformante, y por este camino, al ADN transformante.

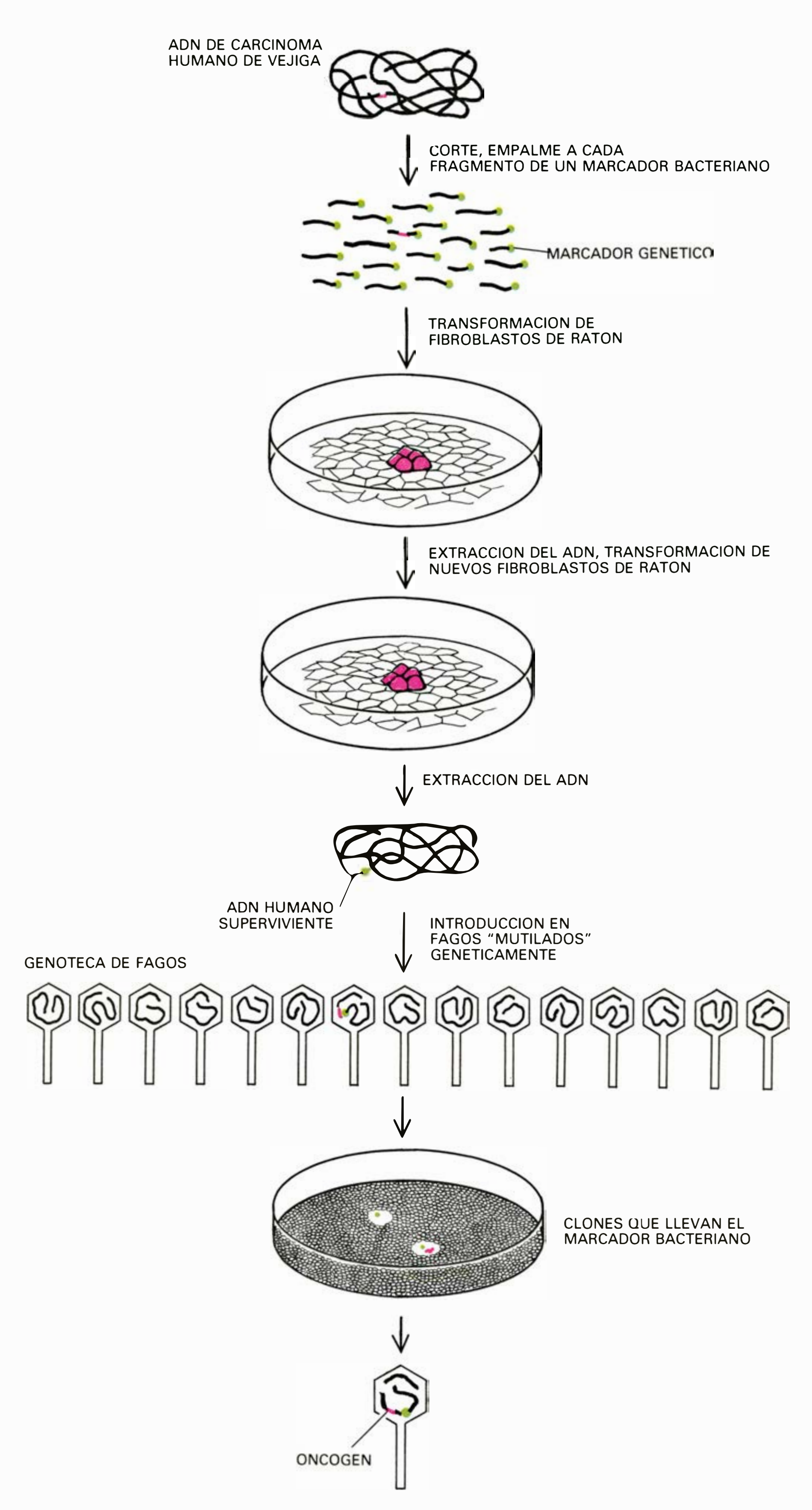
Shih, en mi laboratorio, recurrió a una estrategia diferente. Empezó con ADN de carcinoma de vejiga humano, ácido nucleico con el que transformó células de ratón. Utilizó a continuación el ADN de las células de ratón transformadas para volver a transformar células de ratón, y con el ADN de estas últi-

mas células construyó la genoteca. El ADN empleado para construir la genoteca contenía una cantidad muy pequeña de ADN humano, y precisamente el responsable de la actividad transformante. El resto del ADN humano se había perdido durante las sucesivas transferencias, ya que con éstas sólo se seleccionó el ADN dotado de capacidad transformante. Para buscar este ADN, Shih se sirvió de una sonda que sólo reconoce ADN de origen humano: una sonda específica de secuencias *Alu*. Estas secuencias, muy cortas, sólo se encuentran en ADN humano y están desperdigadas aleatoriamente por todo el genoma. Examinando su genoteca con la sonda específica de *Alu*, Shih encontró y aisló un clon de fagos que llevaba ADN humano. Con un experimento de transferencia génica, demostró que dicho ADN portaba la actividad transformante buscada.

El grupo de Michael Wigler, en el laboratorio de Cold Spring Harbor, utilizó un método completamente diferente. Emplearon ADN de carcinoma de vejiga humano, lo fragmentaron y, a cada uno de los segmentos, le empalmaron cierto marcador genético bacteriano (un gen con actividad conocida y seleccionable). Mediante sucesivas transferencias génicas consiguieron una preparación de ADN de ratón que llevaba sólo una pequeña cantidad de ADN humano, el provisto de capacidad transformante, que portaba junto a él la copia correspondiente del marcador bacteriano. A continuación construyeron la genoteca insertando el ADN de ratón en fagos que tenían defectos en las secuencias determinantes de varias de sus proteínas. Sólo aquellos fagos "mutilados" que incorporasen el marcador bacteriano capaz de corregir su defecto genético podrían multiplicarse bien en un cultivo bacteriano. Entre los pocos fagos viables que alcanzaron a multiplicarse, había varios que llevaban incorporado el segmento transformante.

Oncogenes

Con tres metodologías diferentes y en tres laboratorios distintos, se había logrado aislar el ADN transformante. En todos los casos esta actividad radicaba en segmentos discretos de ADN. Ya podíamos dejar de hablar de aquel vago "principio transformante". Con todos los métodos de clonación molecular utilizados se obtenía un solo segmento de ADN, con un único gen de estructura definida. Estos genes clona-



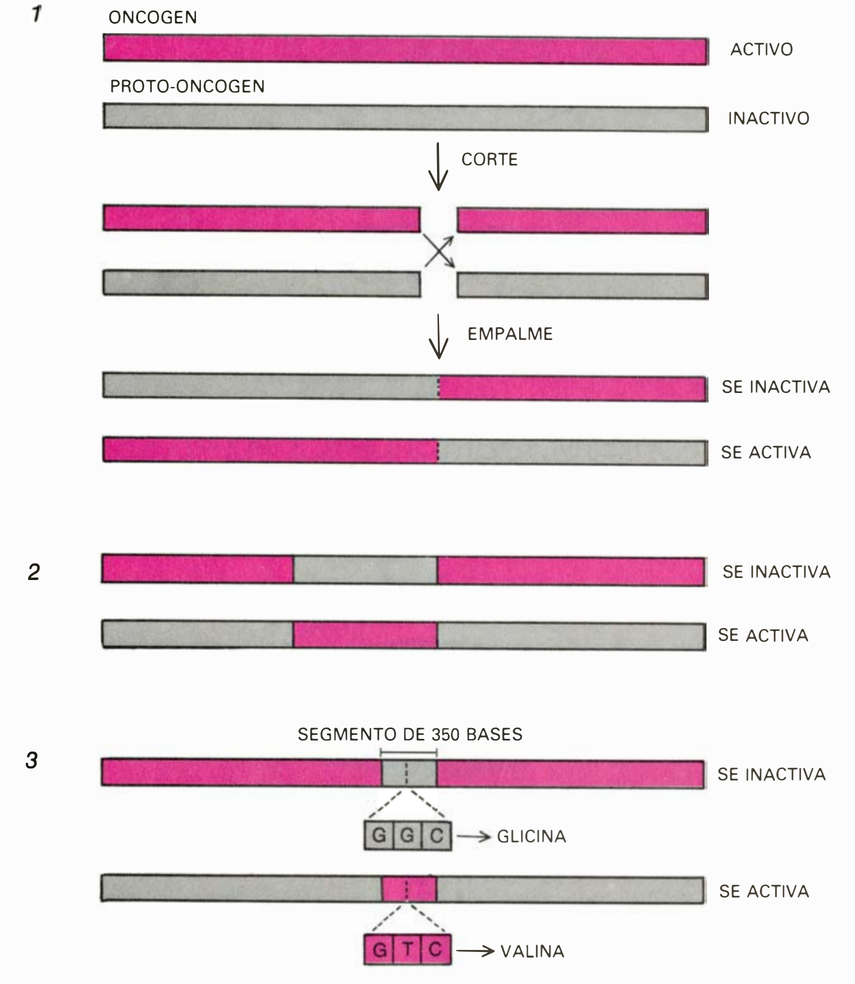
7. UN MARCADOR GENETICO BACTERIANO utilizó como sonda el grupo de Michael Wigler. Fragmentaron el ADN de células de cáncer de vejiga y a cada segmento le empalmaron una copia del marcador bacteriano (*verde*). Tras varias transferencias consiguieron ADN de ratón que no llevaba ADN humano, salvo el gen transformante (*rojo*), unido a una copia del marcador. Construyeron una genoteca introduciendo los fragmentos de ADN de ratón en una estirpe de fago portadora de un defecto que le impedía su desarrollo. Sólo aquellos clones de fagos que llevasen el marcador bacteriano formarían halos sobre un césped de *E. coli*. Algunos de estos clones eran portadores, además, del ADN del cáncer de vejiga.

dos tenían una alta actividad biológica. Con dos microgramos de ADN del carcinoma de vejiga original se obtenía, por término medio, una colonia de células transformadas. Con una cantidad equivalente del gen transformante clonado se obtenían del orden de 50.000 colonias. La actividad transformante, que antes se había atribuido globalmente al ADN de la célula del tumor, podía asignarse ahora a un solo gen: un oncogén, un gen del cáncer.

¿Cómo se origina un oncogén humano? Tres laboratorios dieron con la respuesta: el de Wigler, el mío y el de Mariano Barbacid, del Instituto Nacional del Cáncer de los Estados Unidos. En ello desempeñó un papel fundamental el oncogén de carcinoma de vejiga humano, aislado en la forma descrita anteriormente a partir de una línea celular cancerosa denominada indistintamente *EJ* o *T*24. Cuando se utilizó el oncogén clonado para sondear el genoma humano normal, hibridó intensamente con una secuencia de ADN semejante, presente en dicho ADN normal. Además, se vio que el oncogén y el gen normal tenían el mismo tamaño. Posteriormente se comprobó que ambos genes eran casi idénticos, aunque desde luego sabíamos que no lo eran del todo, pues se comportaban de forma diferente: los clones del oncogén transformaban células, mientras que las del gen normal no tenían tal efecto. No podían ser gemelos idénticos. Por el contrario, el oncogén parecía ser una versión ligeramente alterada del gen normal. Al gen normal se le suele denominar proto-oncogén.

El precursor del oncogén de carcinoma de vejiga no es el único proto-oncogén existente en el ADN humano. Hasta ahora se han encontrado proto-oncogenes correspondientes a otros tres oncogenes activos de tumores humanos. ¿Por qué se mantienen en el genoma humano genes que, con ligeras alteraciones, pueden convertirse en agentes capaces de transformar células y generar tumores? ¿Por qué razón un organismo iba a llevar las semillas de su propia destrucción?

No se conocen aún todas las respuestas a estas importantes cuestiones, aunque una cosa parece obvia: si los proto-oncogenes no desempeñasen un papel vital en el metabolismo celular normal, no se habrían conservado en el genoma. Y el hecho es que se han conservado. En el ADN de ciertos mamíferos, en gallinas e incluso en la mosca de la fruta (*Drosophila*), se han encontrado secuencias similares a los proto-oncogenes humanos. Esto quiere decir que los antecesores de los proto-oncogenes humanos debían haber evolucionado ya hace 600 millones de años, cuando vivía el antecesor común de seres humanos y moscas. Los proto-oncogenes no habrían permanecido inalterables durante tanto tiempo, a menos que fueran, y continúen siendo, indispensables. Se desconoce el papel exacto que corresponde a dichos genes en el metabolismo, aunque hay indicios de que tienen que ver con el control de la proliferación celular. Volveremos sobre ello más adelante.



8. MUTACION PUNTUAL que convierte un proto-oncogén en el oncogén activo de la línea celular *EJ*, del carcinoma de vejiga. Se identificó tras localizar en ambos genes el segmento responsable de la diferencia entre ellos; posteriormente se secuenciaron ambos segmentos. El proto-oncogén y el oncogén (*1*) se cortaron por el mismo sitio y se unieron los segmentos resultantes tal como indica la figura. Mediante transferencia génica se comprobó cuál de los genes recombinantes había perdido su actividad y cuál la había adquirido. El experimento se repetía utilizando cada vez segmentos más pequeños (*2*), hasta que se llegó a una secuencia de sólo 350 nucleótidos, que resultó ser la crítica (*3*). En este segmento apareció una sola mutación: en el proto-oncogén había una guanina (*G*), que se sustituía por timina (*T*) en el oncogén. Así, el triplete del codón correspondiente codificaba el aminoácido valina donde antes había glicina.

Oncogenes víricos

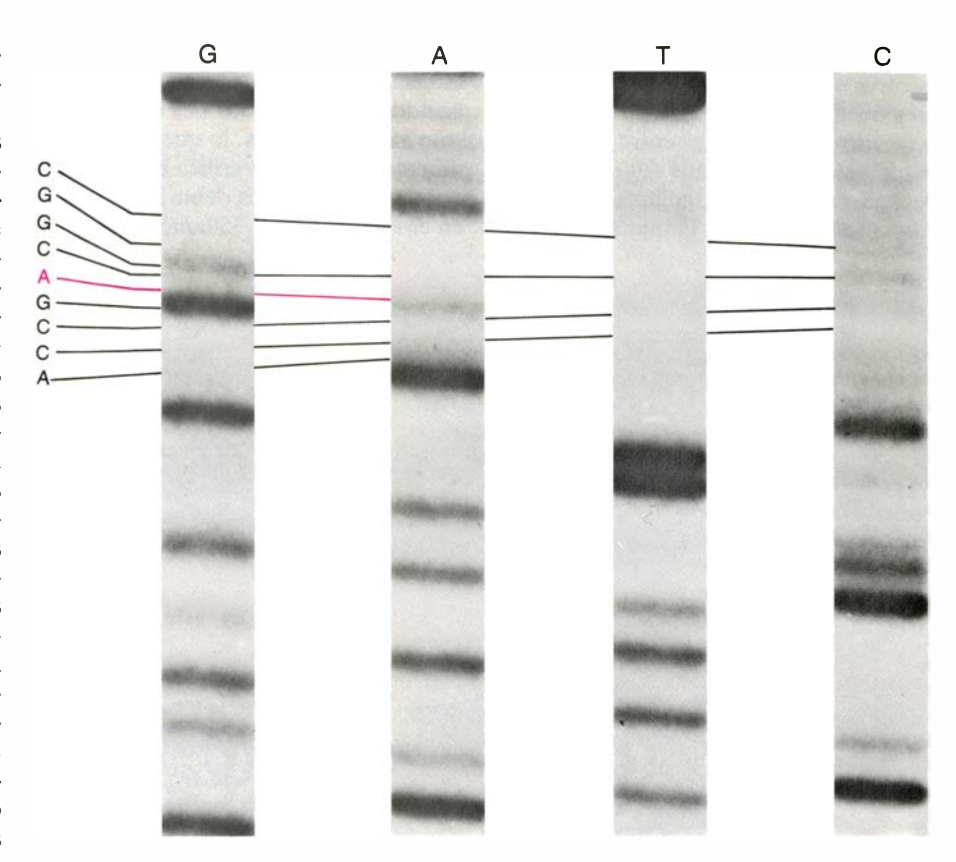
Los experimentos de transferencia de genes permitieron descubrir un tipo de oncogenes y proto-oncogenes. No obstante, se conocía con anterioridad la existencia de otro tipo, gracias a los trabajos con retrovirus, un grupo de virus que provocan cáncer en algunos animales y cuyo material genético no es ADN, sino ARN. Este ARN se "retrotranscribe" a ADN cuando el virus infecta a una célula. Se demostró que los retrovirus llevaban un gen responsable de la inducción del cáncer, un oncogén. J. Michael Bishop y Harold E. Varmus y sus colaboradores, de la Facultad de Medicina de la Universidad de California en San Francisco, demostraron que estos oncogenes no eran realmente víricos, sino que tenían un origen celular. Se trataba de proto-oncogenes "capturados" y transportados (transducidos) por los retrovirus que infectan células animales. Se activan y se convierten en oncogenes una vez incorporados al genoma del retrovirus. Posteriormente son utilizados para transformar las células que ellos mismos infectan. Hasta la fecha se conocen 17 de estos proto-

oncogenes celulares, todos ellos asociados, como oncogén activo, a un determinado retrovirus.

Parecía, pues, que había dos grupos de proto-oncogenes normales y sus oncogenes derivados activos, diferenciados por su forma de activarse. Uno de los grupos, descubierto en los experimentos de transferencia de genes, correspondía a los oncogenes que se originaban por mutación de los proto-oncogenes. El otro, descubierto antes por los virólogos, comprendía aquellos genes que los retrovirus activaban. Sabemos ahora que los dos grupos no son mutuamente excluyentes. Recientes experimentos de hibridación han puesto de manifiesto que algunos oncogenes de tumores humanos son muy parecidos a los oncogenes que llevan ciertos retrovirus que infectan ratas. Por ejemplo, el oncogén de carcinoma de vejiga humano, del que hemos hablado anteriormente, es bastante similar al oncogén que lleva el virus del sarcoma de Harvey, que a su vez procede del genoma de rata. Ello indica que el mismo proto-oncogén puede activarse a través de dos mecanismos independientes: por mutación, en seres humanos, y por la acción de retrovirus, en ratas. Se ha encontrado otro proto-oncogén celular que se comporta de forma similar. Este se puede activar, bien porque se integre en el virus del sarcoma de Kirsten, de ratas, o bien por mutaciones, dando así lugar a los oncogenes presentes en los carcinomas humanos de colon, pulmón, vejiga y páncreas, así como en algunos sarcomas humanos.

Los dos proto-oncogenes celulares, asociados respectivamente con los virus de Harvey y Kirsten, están emparentados evolutivamente. Son miembros de una familia de genes, llamada *ras*, estudiada en detalle por Edward M. Scolnick y sus colaboradores, primero en el Instituto Nacional del Cáncer y actualmente en los Laboratorios Merck de West Point, Pennsylvania. Recientemente, el grupo de Wigler ha relacionado un tercer oncogén humano con la familia *ras*, aunque todavía no se ha encontrado en ningún retrovirus. Este nuevo oncogén, llamado *N-ras*, se encuentra en el ADN de varios tipos de leucemias, un linfoma, un neuroblastoma, un carcinoma de colon y varios sarcomas.

La presencia de los oncogenes *ras* en una amplia gama de tumores indica que la activación de un proto-oncogén celular concreto no está específicamente limitada a un tejido particular. El mismo gen puede activarse en varios tejidos, originando en cada caso un tipo diferente de tumor.



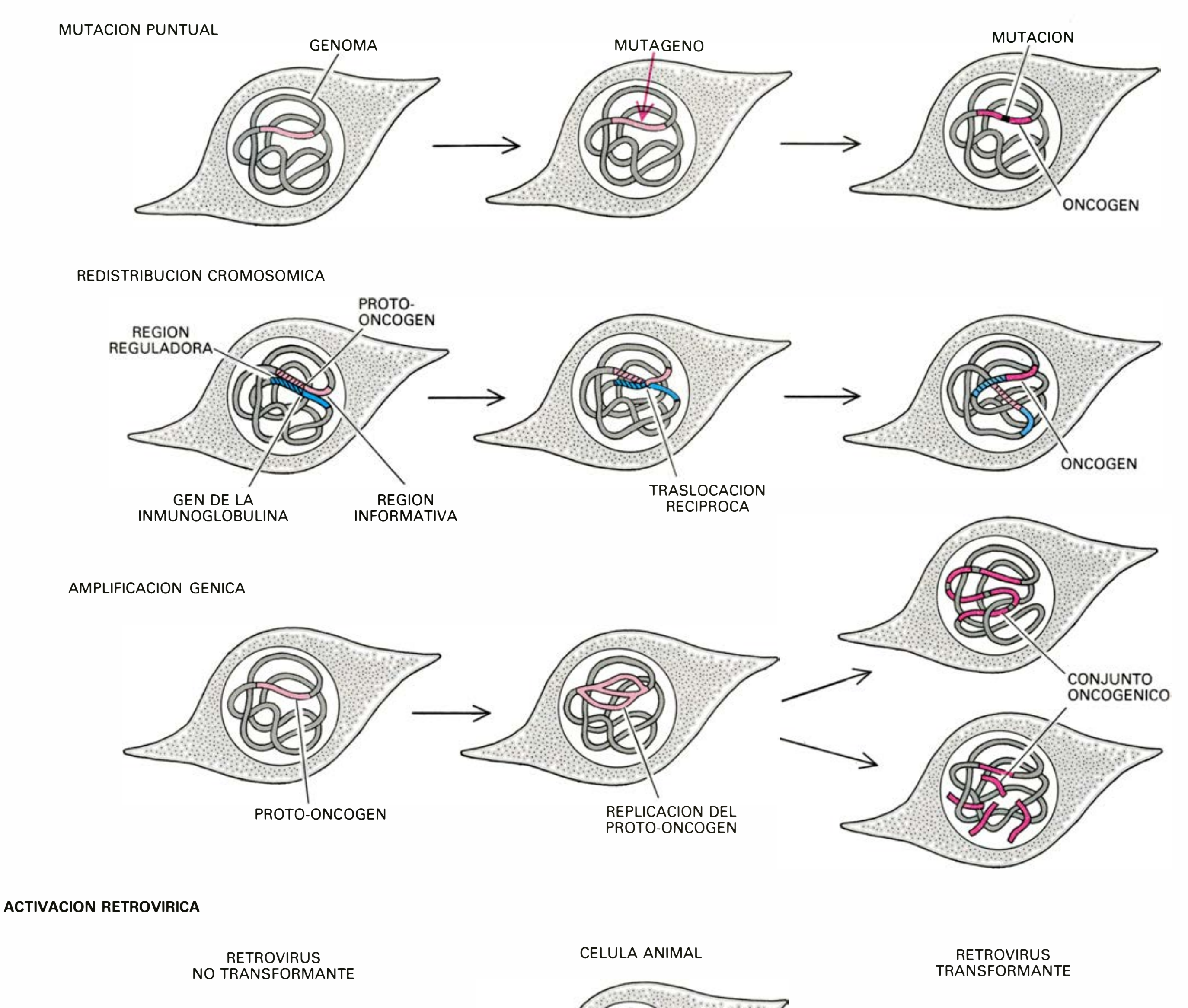
9. SECUENCIA DE BASES del segmento que contenía 350, en el proto-oncogén y en el oncogén. La determinó Ravi Dhar. Esta autorradiografía de un gel muestra la mutación crítica que presenta el oncogén. La distribución escalonada de bandas representa fragmentos cada vez más cortos del segmento (de arriba abajo). Se prepararon cuatro series de fragmentos, que se marcaron por el método de los dideoxinucleótidos, un proceso de secuenciación de ADN que detiene la síntesis del ADN cuando se añade a la cadena uno de los cuatro nucleótidos. Cada una de las series se dispuso en un canal distinto del gel. (Se indica arriba el nucleótido donde se para la síntesis.) Un examen minucioso de la posición que ocupa cada una de las bandas nos permite leer la secuencia. Como este gel se preparó utilizando la cadena del oncogén complementaria a la determinante, la mutación aparece (*color*) como adenina (*A*) en vez de timina (*T*).

Esta simple apreciación es ya de por sí un avance considerable. Desde la perspectiva de un biólogo molecular, el cáncer no constituye ya un conjunto de cien enfermedades, cada una de ellas caracterizada por un tipo de tumor diferente. Por el contrario, empieza a considerarse la existencia de unos pocos mecanismos moleculares comunes a todos los tipos de tumores. No obstante, este punto de vista necesita más apoyo experimental, ya que por ahora son pocos los tumores humanos que se han estudiado en detalle.

¿Qué tipo de mutación?

El paso de proto-oncogén benigno a oncogén activo parecía depender de una mutación; ignorábamos, no obstante, de qué tipo. Una vez que se dispuso de versiones clonadas de ambos genes, se pudo localizar la mutación con precisión. Una forma de hacerlo era determinar la secuencia completa de nucleótidos de ambos genes (con unos 5000 pares de bases cada uno) y compararlas. En la práctica es mejor delimitar primero la región de los genes que presenta el cambio crítico, lo que se puede hacer generando recombinantes formados por parte de los dos genes, remedando con técnicas de ingeniería genética los entrecruzamientos de la genética clásica [*véase la figura 8*]. Tres grupos distintos emprendieron estos experimentos: el de mi laboratorio, el de Barbacid y el de Wigler. En primer lugar, se cortaron los dos genes por el mismo sitio, utilizando restrictasas, que son enzimas que cortan el ADN en los sitios donde se encuentra una secuencia específica de nucleótidos. A continuación, se mezclaron los segmentos procedentes de ambos genes, utilizando el enzima ligasa de ADN para empalmarlos. De esta forma se fabricaban dos genes híbridos, cada uno de los cuales llevaba parte del oncogén y parte del proto-oncogén. Las moléculas híbridas se transfirieron a fibroblastos de ratón para ver cuál de ellas había perdido la actividad oncogénica y cuál la había adquirido.

En mi laboratorio, Clifford J. Tabin y otros colegas consiguieron "cercar" la región que nos interesaba, tras sucesivos entrecruzamientos entre segmentos cada vez más pequeños del oncogén de vejiga y su correspondiente segmento del proto-oncogén. Finalmente lograron convertir el proto-oncogén en una molécula transformante activa introduciéndole un segmento del oncogén de sólo 350 nucleótidos, lo que venía a significar que la lesión crítica que diferenciaba a ambos genes debía encontrarse en ese segmento de 350 nucleótidos. El análisis de la secuencia de las versiones activa e inactiva del segmento, realizado por Ravi Dhar, del Instituto Nacional del Cáncer, llevó a un espectacular e inesperado descubrimiento. La secuencia de 350 nucleótidos del oncogén y la correspondiente del proto-oncogén



10. FORMAS DE ACTIVAR UN PROTO-ONCOGEN; hay cuatro, tres de las cuales implican alteración del ADN celular y la cuarta la intervención de un retrovirus. En una de las formas no víricas se trata de una mutación puntual. Si un proto-oncogén de una célula normal sufre una mutación por acción de radiaciones o un carcinógeno químico, la mutación puede hacer que cambie la información que determina la síntesis de una proteína. La proteína alterada puede inducir uno o más de los cambios asociados con el crecimiento canceroso. Otra posibilidad: la región determinante de la proteína de un proto-oncogén se recombina, en el curso de reorganizaciones cromosómicas, con la región reguladora de un gen de inmunoglobulinas, haciendo que la información genética del proto-oncogén se exprese a un nivel desacostumbrado. Una tercera forma de activación es la multiplicación (amplificación): en virtud de una alta tasa de replicación, dentro de una célula puede llegar a haber muchas copias de un proto-oncogén, secuencialmente repetidas a lo largo de un cromosoma, o como partículas extracromosómicas dispersas. También en este caso la consecuencia es un exceso de expresión. Los retrovirus, que producen cáncer en muchos animales, adquieren sus oncogenes al infectar una célula animal. Durante la infección, el virus "captura" una secuencia celular que incluye el proto-oncogén y la integra en su propio genoma. En este caso, el proto-oncogén puede activarse por asociación con una región reguladora del genoma vírico, tal como aquí se muestra, o bien porque sufra una mutación mientras el retrovirus lo transporta (transduce).

diferían en una sola base: una guanina del proto-oncogén aparecía sustituida por una timina en el oncogén. Por tanto, el cambio de un solo nucleótido, una "mutación puntual", en un gen humano normal que contiene alrededor de 5000 nucleótidos puede convertirlo en un oncogén. Por primera vez se definía con precisión una lesión genética que podía llevar al desarrollo de un tumor. Quizás un paso crítico en la iniciación de muchos cánceres sea la mutación de un proto-oncogén, instada por alguna "agresión" ambiental a su ADN.

Recientemente se han descrito otros mecanismos de activación de proto-oncogenes. Uno de ellos, consistente en una reorganización cromosómica, fue descubierto por Grace L. C. Shen-Ong y Michael D. Cole, del Hospital Clínico de la Universidad de St. Louis, y por Rebecca A. Taub y sus colegas, en el laboratorio de Philip Leder, de la Facultad de Medicina de Harvard. Este mecanismo es responsable de la activación de un proto-oncogén llamado *myc*, cuya versión con actividad oncogénica se ha descubierto en mielomas de ratón y en un cáncer humano, el linfoma de Burkitt. El proto-oncogén celular se separa de su posición habitual, en el extremo de un cromosoma, traslocándose al extremo de otro cromosoma distinto. En este segundo cromosoma, el gen traslocado se inserta junto a genes responsables de la síntesis de inmunoglobulinas (anticuerpos). Durante la respuesta inmunológica a un antígeno, tales genes se transcriben muy activamente. Aparentemente, la asociación del proto-oncogén con un gen de inmunoglobulinas trastorna la regulación del proto-oncogén y le impone una nueva función biológica.

En otros casos se han hallado numerosas copias de un proto-oncogén en cada célula, en vez de las dos copias normales, características de la mayoría de los genes celulares. Esta multiplicación (amplificación) génica, y el alto grado de expresión génica que lleva consigo, también parece tener un efecto oncogénico. En otras palabras, existen ciertos mecanismos moleculares que pueden forzar a un gen normal a convertirse en un oncogén. En el caso de los genes de tumores de mamíferos, dentro de estos mecanismos se incluyen las mutaciones puntuales, reorganizaciones cromosómicas y multiplicación génica. En el caso de los genes asociados con retrovirus, el gen trasducido puede sufrir una mutación o puede insertarse junto a una región reguladora vírica que incremente su nivel de expresión.

Proteínas del cáncer

La búsqueda de los mecanismos responsables del cáncer ha progresado bastante, como lo demuestra la transición ocurrida desde que se asignaba la responsabilidad a unos oscuros "factores genéticos" hasta adjudicársela a moléculas de ADN y finalmente a genes concretos dentro de esas moléculas. La búsqueda parecía acabar con el descubrimiento de una alteración molecular, definida además con precisión, en uno de tales genes. Lo cierto es que están todavía por resolver los problemas más difíciles.

Aunque se da por supuesto que la creación de un oncogén puede originar un cáncer, no tenemos aún idea de cómo el gen alterado contribuye a este proceso. ¿Qué función desempeña un oncogén? Al igual que otros genes, determina la estructura de una proteína. El grupo de Scolnick ha demostrado que muchos de los genes *ras* determinan la síntesis de una proteína de un peso molecular de 21.000. La mutación puntual observada en el oncogén del carcinoma de vejiga altera la estructura de la proteína. Cada uno de los aminoácidos que componen una proteína está especificado por un codón de tres bases en el gen que determina la estructura de la proteína. La mutación puntual en cuestión provocaba el cambio de un codón *GGC* del proto-oncogén por *GTC* en el oncogén. Ello produce un cambio en la estructura de la proteína del oncogén, que ahora lleva el aminoácido valina en un sitio normalmente ocupado por una glicina. Por lo que se ve, esa única sustitución obliga a la proteína a asumir nuevas funciones, ocasionando una alteración profunda del metabolismo celular.

En el núcleo mismo del proceso canceroso deben estar implicados elementos celulares tales como estas proteínas de los oncogenes, lo que desconocemos es qué función desempeñan éstas. Podemos suponer que están implicadas en la regulación del crecimiento celular, crecimiento que es normal cuando se halla presente la proteína determinada por un proto-oncogén, mientras que si aparece la versión del oncogén, se induce el crecimiento canceroso.

Cuando sepamos con precisión qué función desempeñan las proteínas de los oncogenes se podrán buscar antagonistas que inhiban tal función, y quizá se logre desarrollar así una terapia específica contra un defecto conocido y fundamental de la célula cancerosa. Sin embargo, es probable que este defecto no sea el único, ya que la carcinogénesis parece ser un proceso en el que intervienen muchos factores. Diferentes indicios apuntan a que una célula normal debe sufrir varias alteraciones independientes antes de convertirse en una "buena" célula tumoral. Una mutación puntual, que origina una proteína defectuosa, constituye sólo un paso, que afecta a un único gen. Presumiblemente, la creación de un oncogén sea una de las etapas en la formación de un tumor, pero hay otras. El oncogén puede ser necesario, pero difícilmente suficiente. Empiezan a recabarse datos relacionados con la naturaleza de esas otras etapas.

Se han encontrado ciertos tumores en los que se han originado dos oncogenes como consecuencia de fenómenos de activación independientes. Por ejemplo, William S. Hayward, de la Universidad Rockefeller, y Susan M. Astrin, del Instituto para la Investigación del Cáncer, han observado que el proto-oncogén *myc* se activa en linfomas de gallina, mientras que el grupo de Cooper ha encontrado un segundo proto-oncogén, *lym-B*, que también se activa en las mismas células tumorales. En el linfoma de Burkitt y en una leucemia promielocítica humana se han descubierto dos oncogenes en estado activado. Todos estos datos sugieren que la compleja naturaleza de la carcinogénesis puede deberse, en parte, a la existencia de varios oncogenes distintos. Cada oncogén podría provocar una alteración diferente en la célula, y todos juntos colaborarían en la expresión de muchas de las características del fenotipo tumoral.

He planteado sólo dos de los numerosos problemas que quedan por resolver. Resulta alentador, no obstante, que los resultados obtenidos en varias líneas de investigación empiecen a darle sentido a una enfermedad que hace sólo cinco años parecía incomprensible. Los recientes descubrimientos en el nivel génico son coherentes con las primeras ideas sobre la carcinogénesis, basadas en datos epidemiológicos y en los estudios sobre transformación.

Se dispone ahora de poderosas herramientas que permitirán avanzar más deprisa en los próximos años. Si no se tuerce el rumbo, a finales de esta década conoceremos con detalle los mecanismos moleculares que determinan el cáncer.

Superestructuras de estado sólido

Existen cristales constituidos por materiales semiconductores dispuestos en forma de capas cuyo espesor se mide en átomos. Las propiedades de los mismos son muy prometedoras para la física del estado sólido y la electrónica

Gottfried H. Döhler

"Tú ves las cosas y te preguntas ¿por qué? En cambio, yo me imagino las cosas que nunca han existido y me pregunto ¿por qué no?" Estas palabras de George Bernard Shaw reflejan el espíritu según el cual algunos físicos especialistas en estado sólido empezaron a especular hace una década acerca de las propiedades y la utilidad de ciertos semiconductores que en aquel entonces no existían todavía. Estos físicos fueron lo bastante realistas como para enfocar sus ideas hacia los semiconductores que podrían crearse en principio, aunque la naturaleza no los hubiera producido. En concreto, tales investigadores previeron la existencia de ciertas superestructuras de estado sólido, constituidas por unos semiconductores especiales que se obtienen a partir de dos materiales con propiedades electrónicas diferentes, depositados en forma de capas delgadas alternadas, o introduciendo algunas impurezas en determinados materiales semiconductores.

Todas estas esperanzas se vieron coronadas por el éxito en un doble sentido. En primer lugar, se demostró que dichas superestructuras podían "crecer" con una gran precisión. En segundo lugar, las propiedades de estas superestructuras confirmaron las previsiones de los investigadores. Algunas propiedades son sorprendentes: por ejemplo, disminución paradójica de la corriente eléctrica al aumentar el voltaje o persistencia de electrones libres en la superestructura durante horas, y no ya sólo millonésimas de segundo. La predicción de estas propiedades incluso antes de conocerse dichas superestructuras nos habla del progreso experimentado por la física del estado sólido. Esta circunstancia apunta también hacia una tendencia cuya importancia va creciendo de día en día: la utilización de la teoría cuántica de los sólidos como guía para la obtención de nuevos materiales.

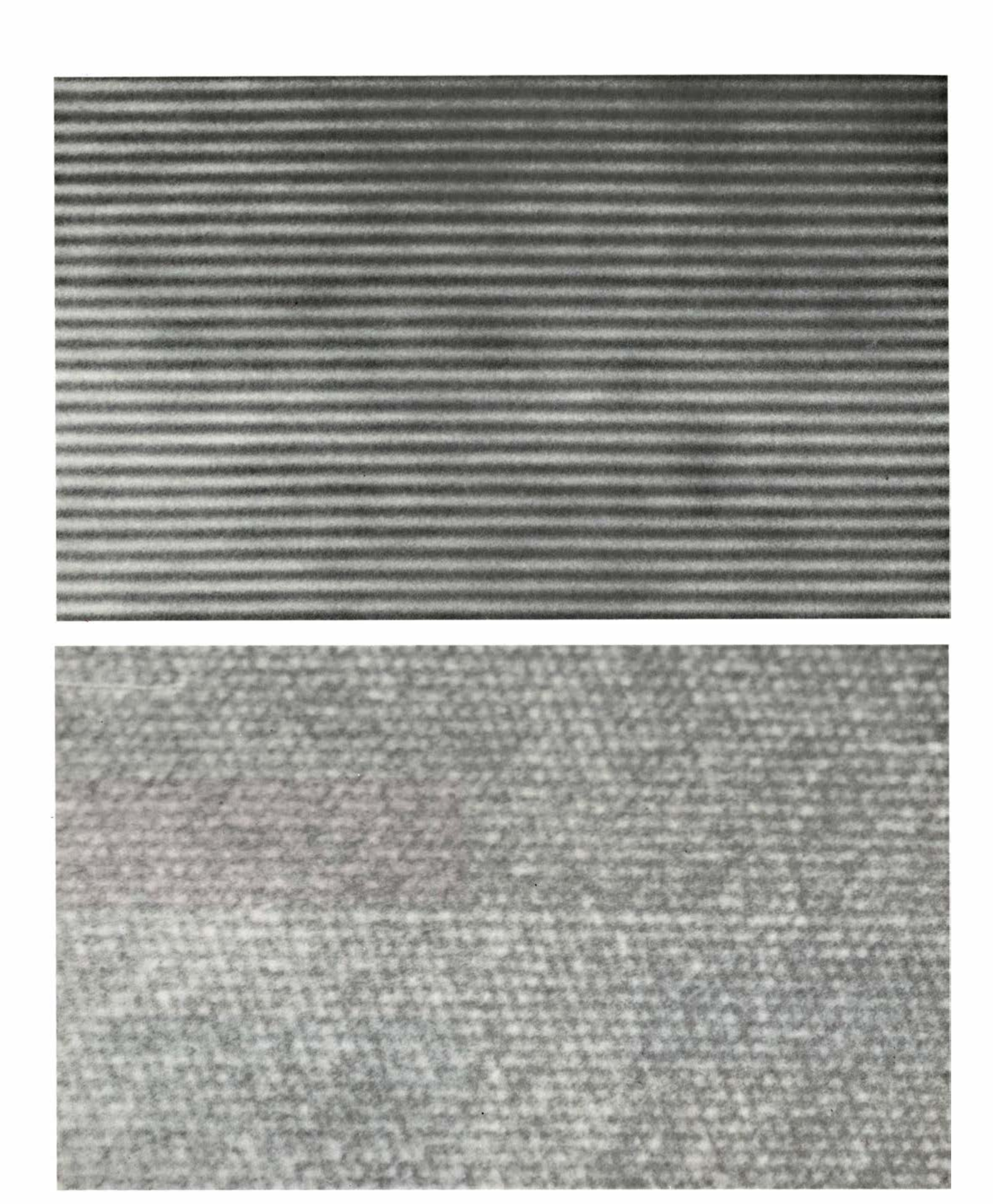
Antes de describir las propiedades de las superestructuras de cristales sólidos conviene recordar las características fundamentales de los semiconductores. Consideremos un cristal puro y perfecto de un material semiconductor, el silicio por ejemplo. A la temperatura del cero absoluto (cero grados Kelvin) este material se convertiría en un aislante perfecto, pues sus electrones (los portadores de carga responsables de la conducción eléctrica en los metales) serían incapaces de moverse. La explicación de esta falta de movilidad radica en que los electrones de un semiconductor sólo pueden ocupar ciertos niveles energéticos, niveles que forman "bandas" de energía. Un cierto intervalo energético separa la "banda de valencia", completamente ocupada por los electrones que constituyen los enlaces químicos que mantienen unidos a los átomos del cristal, de la "banda de conducción", que está vacía. A la temperatura de cero grados, los electrones no pueden remontar dicho intervalo. Sin embargo, a temperaturas más elevadas, un electrón puede adquirir ocasionalmente la energía necesaria para alcanzar la banda de conducción. En otras palabras, ese electrón se puede mover a lo largo del cristal.

¿Qué entendemos exactamente por "ocupación de una banda"? Supongamos que en un sólido existen *N* átomos ligados. Si estuviera aislado, cada uno de estos átomos poseería cierto número de estados energéticos disponibles para sus electrones. Según el principio de exclusión de Pauli, que rige en física cuántica, cada estado sólo puede aceptar a lo sumo dos electrones y los espines (momentos angulares intrínsecos) de estos dos electrones deben estar alineados en direcciones opuestas. Ahora bien, en un sólido estas condiciones se alteran: cada estado se desdobla en un total de 2*N* subestados. En un sólido de tamaño apreciable, el valor de *N* es al menos de 10^{23}. En consecuencia, los estados son tan numerosos y sus energías tan poco espaciadas entre sí que prácticamente forman un continuo. En cada átomo, sin embargo, cada uno de estos estados sólo puede aceptar dos electrones.

Supongamos ahora que cada átomo posee un número par de electrones. En estos átomos los estados de energías más bajas están totalmente ocupados (con dos electrones en cada estado); por tanto, la banda de valencia está llena. Supongamos además que la banda de valencia y la banda de conducción no se solapan entre sí, aunque sus energías no difieran excesivamente. En tal caso, el sólido es un semiconductor. (Si la banda de valencia y la banda de conducción se solaparan, los electrones de los niveles superiores de la banda de valencia "fluirían" libremente hacia los niveles más bajos de la banda de conducción.) Imaginemos que el número de átomos es impar, de forma tal que el estado más alto de la banda de valencia sólo esté parcialmente lleno: en ese caso, el sólido es un metal ordinario.

De esta sumaria exposición se desprende al menos la conclusión de por qué un metal ordinario es un buen conductor. Los electrones del metal se aceleran fácilmente mediante la aplicación de un campo eléctrico, dado que la diferencia entre las energías de un subestado ocupado y un subestado vacío en un cuasi continuo de 10^{23} subestados resulta despreciable, por lo pequeña. Por contra, los electrones de un semiconductor se ven obligados a cruzar el intervalo que separa la banda de valencia y la banda de conducción antes de que puedan alcanzar a esta última. El tránsito se puede producir de muy diversas maneras, por ejemplo absorbiendo el semiconductor un fotón (un cuanto de luz) con una energía mayor que la anchura de dicho intervalo.

Todavía existe otra característica



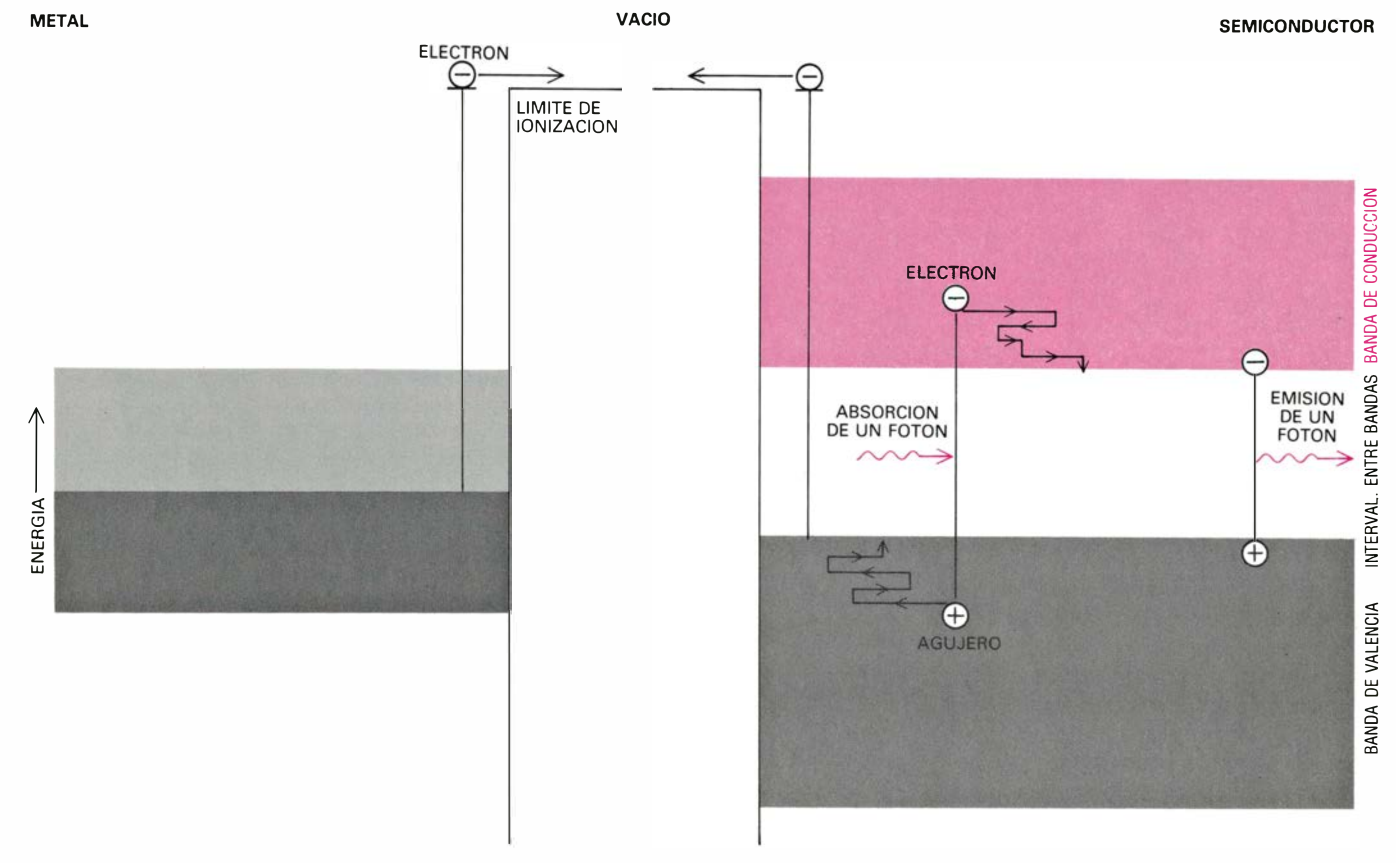
1. SUPERESTRUCTURA COMPOSICIONAL, constituida por un conjunto periódico de capas alternas de dos semiconductores diferentes. En la microfotografía electrónica que se reproduce en la parte superior de esta figura los semiconductores son el arseniuro de galio y el arseniuro de aluminio. Cada banda oscura corresponde a seis planos atómicos de arseniuro de galio; cada banda brillante, a cuatro planos atómicos de arseniuro de aluminio. La distancia entre dos planos atómicos consecutivos es de $2{,}83 \times 10^{-10}$ metros. En la microfotografía electrónica de la parte inferior se observa la presencia de una superficie de separación simple entre una capa de arseniuro de galio (*mitad superior*) y una capa de arseniuro de galio y aluminio (*mitad inferior*). La superficie en cuestión está situada en el centro de la imagen y está constituida por un plano cuyo espesor es de un solo átomo. La distancia entre los planos es de $3{,}6 \times 10^{-10}$ metros. Cada punto brillante de la imagen está constituido por una columna de átomos vistos desde la parte superior de la superestructura cristalina. Las microfotografías electrónicas han sido obtenidas por Pierre M. Petroff y Arthur C. Gossard en los Laboratorios Bell.

fundamental de los semiconductores. Todo electrón excitado a la banda de conducción de un semiconductor origina la creación de un "agujero" en la banda de valencia, es decir, produce la ausencia de un electrón en el correspondiente átomo. Esta ausencia puede desplazarse de un átomo a otro; dicho agujero se comporta como un portador de carga y contribuye a la conductividad del cristal. (Obsérvese que, en cierto sentido, los términos banda de conducción y banda de valencia pueden inducir a confusión.) La carga de un electrón es negativa; en cambio, la de un agujero es positiva.

Supongamos que la absorción de un fotón cuya energía es superior a la anchura del intervalo que separa las bandas de conducción y de valencia en un semiconductor insta la aparición de un electrón y de un agujero. Estos elementos perderán rápidamente su energía en exceso y calentarán algo el cristal. El electrón "desciende" hasta el extremo de energía más baja de la banda de conducción. Por otra parte, se puede suponer que el agujero "asciende" hasta el extremo superior de la banda de valencia. Si ambos elementos se recombinan, se producirá la emisión de un fotón cuya energía es igual a la anchura de la separación entre las bandas del semiconductor. Resulta evidente que la energía del extremo inferior de la banda de conducción y la energía del extremo superior de la banda de valencia son dos magnitudes importantes en la descripción de un semiconductor.

La primera superestructura de estado sólido que se hizo realidad fue una estructura de tipo composicional o heteroestructura. El nuevo material semiconductor había sido previsto en el año 1970 por Leo Esaki y Ray Tsu, del Centro de Investigación Thomas J. Watson de la International Business Machines Corporation. Esaki, Tsu y sus colaboradores obtuvieron dicho material años más tarde.

Las superestructuras a las que nos estamos refiriendo constan de una serie de capas ultradelgadas de dos semiconductores diferentes alternados. El espesor de cada capa es de sólo algunos centenares de átomos. Los semiconductores se eligen de forma tal que la separación entre sus bandas sea diferente; así, la variación espacial de las propiedades electrónicas del material resultante es mucho más compleja que la que corresponde a un semiconductor ordinario. La forma más directa de que se dispone para determinar el correspondiente esquema de niveles consiste en abordar la superestructura en cuestión como un conjunto de capas de semiconductores, cada una de las cuales presenta las propiedades que le son características. Los cálculos más refinados demuestran que esta sencilla estrategia proporciona unos resultados suficientemente satisfactorios, incluso en el caso de que las capas en cuestión sean de extrema delgadez. Se puede afirmar que la alternancia periódica de las capas origina una alternancia periódica del potencial eléctrico. Concretamente, cada una de las capas del semiconductor cuya separación entre las bandas sea menos ancha crea lo que se conoce por pozo de potencial [*véase la figura 3*]. En el interior de uno de estos pozos de potencial los electrones de la banda de conducción sólo disponen de ciertos niveles energéticos. (Una vez más, cabe recordar aquí que cada uno de estos estados se desdobla en un cuasi continuo o minibanda). Esta situación

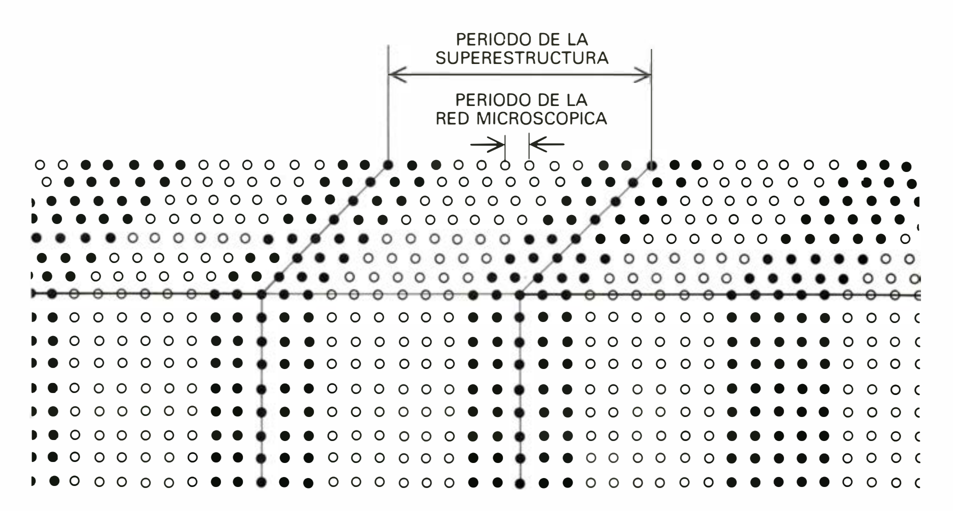


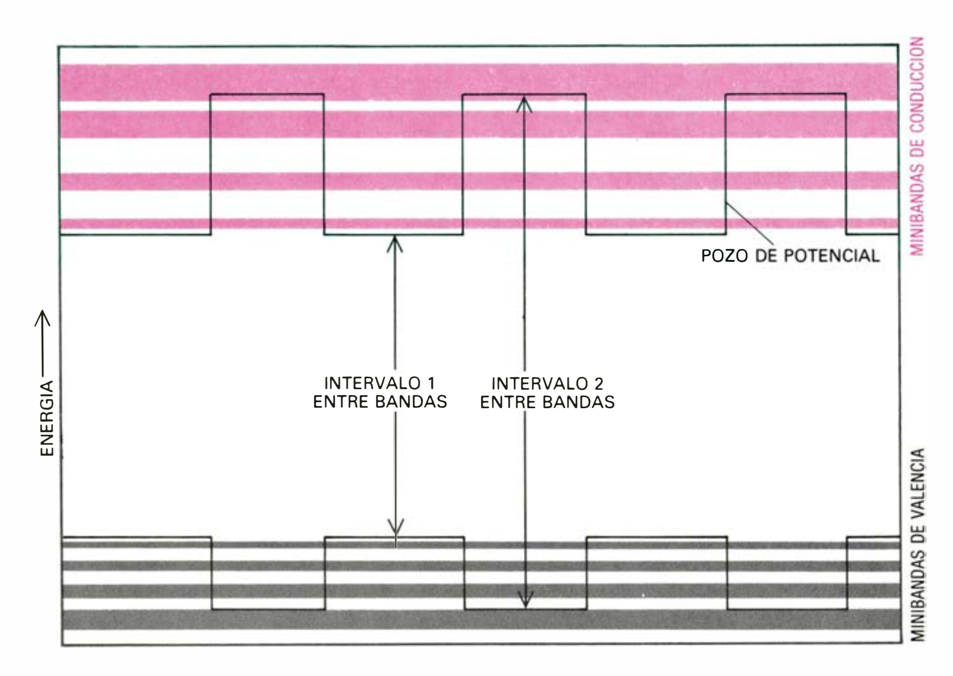
2. PROPIEDADES ELECTRONICAS DISTINTIVAS de un metal y de un semiconductor, representadas en dos diagramas en los que se muestran los niveles energéticos de que dispone cada uno de estos materiales para los portadores de cargas eléctricas. En el caso del metal (*izquierda*) los portadores son electrones, que se mueven libremente en una banda parcialmente llena de energías disponibles. En el caso del semiconductor (*derecha*) los electrones que forman los enlaces entre los átomos llenan una "banda de valencia". Estos electrones se pueden excitar y alcanzan una "banda de conducción", que en otro caso se encuentra vacía, gracias a la absorción de fotones (cuantos de luz) cuya energía es superior a la separación entre bandas. La excitación origina la aparición de "agujeros" en la banda de valencia. Cuando los electrones y los agujeros se recombinan, su energía se disipa por emisión de fotones.

es similar a la que existe en un sólido cristalino, en el que la periodicidad se establece en virtud de la presencia de átomos a intervalos regulares. Cada átomo crea un pozo de potencial y en cada átomo sólo algunos niveles energéticos están disponibles para los electrones.

Sin embargo, entre ambos sistemas se pone de manifiesto una diferencia importante. Las propiedades electrónicas de los átomos vienen predeterminadas por la naturaleza. En cambio, las propiedades electrónicas de las superestructuras se pueden configurar libremente. En primer lugar, los valores de los niveles energéticos que están disponibles para los electrones se pueden modificar mediante una elección adecuada de los semiconductores y del espesor de sus capas. (Básicamente, esto es aplicable también para los niveles disponibles para los agujeros.) Por otra parte, la anchura de las bandas también se puede modificar. Esta anchura viene determinada por la intensidad de la interacción entre los pozos de potencial próximos y aumenta a medida que disminuye el espesor de las capas del semiconductor cuya separación entre las bandas es mayor.

En la práctica, el semiconductor con el intervalo menor suele ser el arseniuro de galio (GaAs); el semiconductor con el intervalo mayor, el arseniuro de galio y aluminio (AlGaAs). (En principio, cabe cualquier relación entre los átomos de galio y aluminio, pero el número total de átomos de galio y aluminio debe ser igual al de átomos de arsénico.) Los cálculos realizados en el caso de una superestructura constituida por una serie de capas alternadas de estos dos materiales demuestran que las correspondientes minibandas son mucho más estrechas que las bandas de un semiconductor ordinario y que dichas minibandas están separadas entre sí por unos intervalos relativamente grandes. El único requisito que se exige para ello consiste en que el espesor de las dos capas se sitúe en el orden de cuatro a diez nanómetros. (Un nanómetro es una millonésima de milímetro.)





3. LAS PROPIEDADES DE UNA SUPERESTRUCTURA COMPOSICIONAL se deducen de la alternancia periódica de los dos semiconductores diferentes que la constituyen. Estos dos semiconductores se escogen de forma tal que la separación entre las bandas de conducción y de valencia tenga anchuras diferentes en cada uno de ellos y que dichas diferencias originen pozos de potencial, es decir, caídas muy acusadas de la energía potencial en el interior del cristal (*dibujo inferior*). Por su parte, los pozos de potencial dividen a la banda de valencia (*gris*) y a la banda de conducción (*color*) en numerosas minibandas. Ciertas propiedades del cristal se pueden modificar a voluntad (anchura y energía de las minibandas) eligiendo bien los semiconductores que lo constituyen y los espesores de las correspondientes capas.

Las minibandas a las que acabamos de referirnos proporcionan las condiciones ideales para que se produzca un fenómeno peculiar de la física del estado sólido conocido por oscilaciones de Bloch, cuya detección parece ser imposible en los sólidos existentes en la naturaleza. Este fenómeno consiste en lo siguiente: cuando un semiconductor se somete a la acción de un campo eléctrico intenso, sus bandas se inclinan. La pendiente de esta inclinación es igual a e veces el voltaje aplicado al cristal dividido por la longitud de dicho cristal (donde e designa la carga del electrón); por tanto, la pendiente aumenta al subir el voltaje. En un semiconductor convencional, los electrones de la banda de conducción son transportados hasta el extremo superior de la banda inclinada [*véase la figura 4*]. Sin embargo, mucho antes de que alcancen dicho extremo, pierden la energía que han ganado del campo eléctrico aplicado y se produce una emisión de "fonones", es decir, se originan vibraciones térmicas en el cristal.

En el caso de las superestructuras a las que nos estamos refiriendo la situación cambia. Las minibandas son más estrechas y, por consiguiente, la probabilidad de que un electrón alcance el extremo superior de una minibanda inclinada es elevada. Los electrones se reflejan en el extremo superior de la minibanda, al no poder atravesar la misma para llegar luego a la minibanda siguiente. Después de sufrir esta reflexión (un fenómeno cuántico conocido como reflexión de Bragg) son devueltos al extremo inferior de la minibanda en cuestión. De hecho, los electrones pueden ser reflejados repetidamente entre los extremos superior e inferior de

dicha minibanda, llevando a cabo muchas de estas oscilaciones de Bloch antes de que se produzca la emisión de un fotón y "caigan" a una energía inferior.

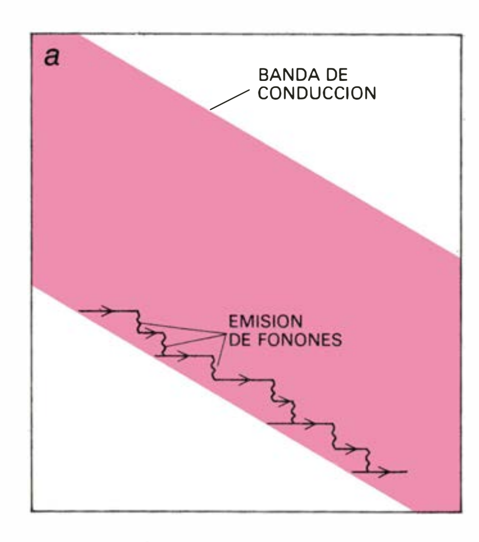
El desplazamiento de la posición media de un electrón (el centro de masas) en la banda de conducción, desplazamiento que tiene lugar en virtud de la emisión de un fonón, disminuye al aumentar la inclinación de dicha banda. En consecuencia, al aumentar el potencial aplicado a una superestructura se puede producir el efecto curioso de que disminuya la intensidad de la corriente que la atraviesa. En otras palabras, el cristal puede presentar una resistencia negativa: este material quizá no consuma energía como sucede en las resistencias, sino que la aporte al circuito oscilante del que forma parte. Por tanto, una superestructura podría actuar como el elemento activo de un generador de ondas electromagnéticas. Ahora bien, al contrario de lo que sucede con otros dispositivos de resistencia negativa, este material reaccionaría de forma casi instantánea a la aplicación de un voltaje y, por ende, generaría una radiación de microondas con longitudes de onda inferiores al milímetro. No se dispone todavía de generadores eficientes para tales longitudes.

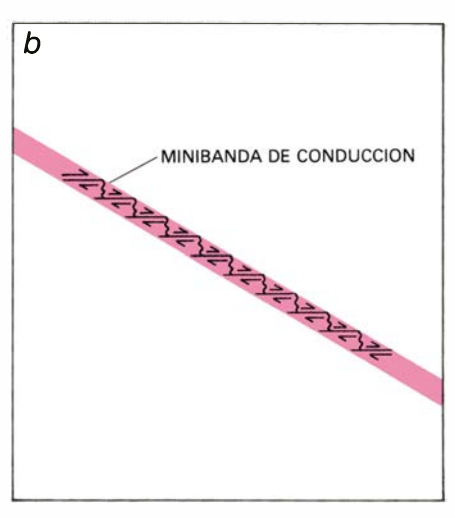
Interesa destacar que Esaki y Tsu predijeron la ocurrencia de las oscilaciones de Bloch en las superestructuras, incluso antes de que éstas existieran. La preparación de dichas superestructuras se llevó a cabo utilizando la técnica de los haces moleculares. Esta técnica fue desarrollada por Alfred Y. Cho y John R. Arthur, Jr., en los Laboratorios Bell, a finales de la década de los años sesenta. El crecimiento de los cristales se consigue haciendo incidir unos haces de los átomos o las moléculas que los constituyen sobre un soporte adecuado (arseniuro de galio). Los haces depositan estos materiales con una velocidad tal que el espesor de la capa que se forma aumenta a razón de un átomo por segundo. Algunos años más tarde, Esaki y L. L. Chang anunciaron la primera comprobación experimental de la predicción que había abierto las puertas a todo este campo de trabajo.

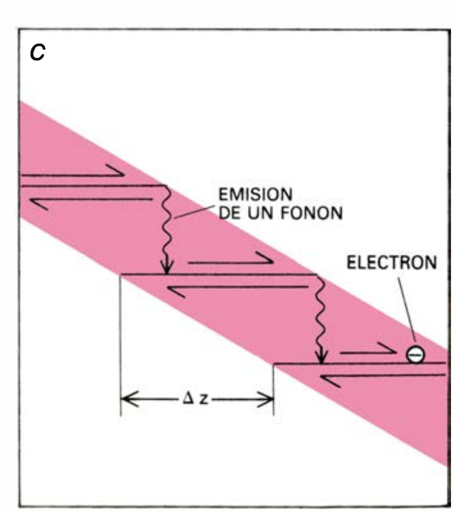
Este primer éxito fue seguido poco después por otro. Raymond Dingle, William Wiegmann, Arthur C. Gossard y Charles H. Henry, de los Laboratorios Bell, habían estado investigando la absorción de la luz por parte de una superestructura constituida por una mezcla de arseniuro de galio y aluminio. Predijeron que los fotones sólo se absorberían si sus energías igualaban o superaban cierto valor umbral que viene determinado, además de por la anchura nominal de la separación entre las bandas del semiconductor, por la anchura efectiva de dicho intervalo, es decir, por la diferencia entre la energía de la minibanda de conducción más baja y la energía de la minibanda de valencia más alta [*véase la figura 3*]. Por otro lado, estos mismos autores predijeron que la absorción aumentaría fuertemente cuando el valor de las energías de los fotones se aproximara al de las diferencias entre las energías de otras parejas de minibandas.

Los ensayos confirmaron todas estas predicciones. Una vez más, quedó plenamente demostrada la validez de la física cuántica en su aplicación a un material producido por el hombre. Por otro lado, dicha confirmación puso de manifiesto que el crecimiento de la superestructura en cuestión se había producido con la precisión prevista. Este hecho planteó además la posibilidad de utilizar estas superestructuras en el estudio de ciertos fenómenos mecánico-cuánticos complejos. En particular, las interacciones entre las partículas de un sistema mecánico-cuántico de muchos cuerpos podrían investigarse bajo condiciones sencillas y bien definidas en el caso de que un pozo de potencial rectangular, practicado con gran precisión en una de estas superestructuras, estuviera ocupado por muchos portadores de carga.

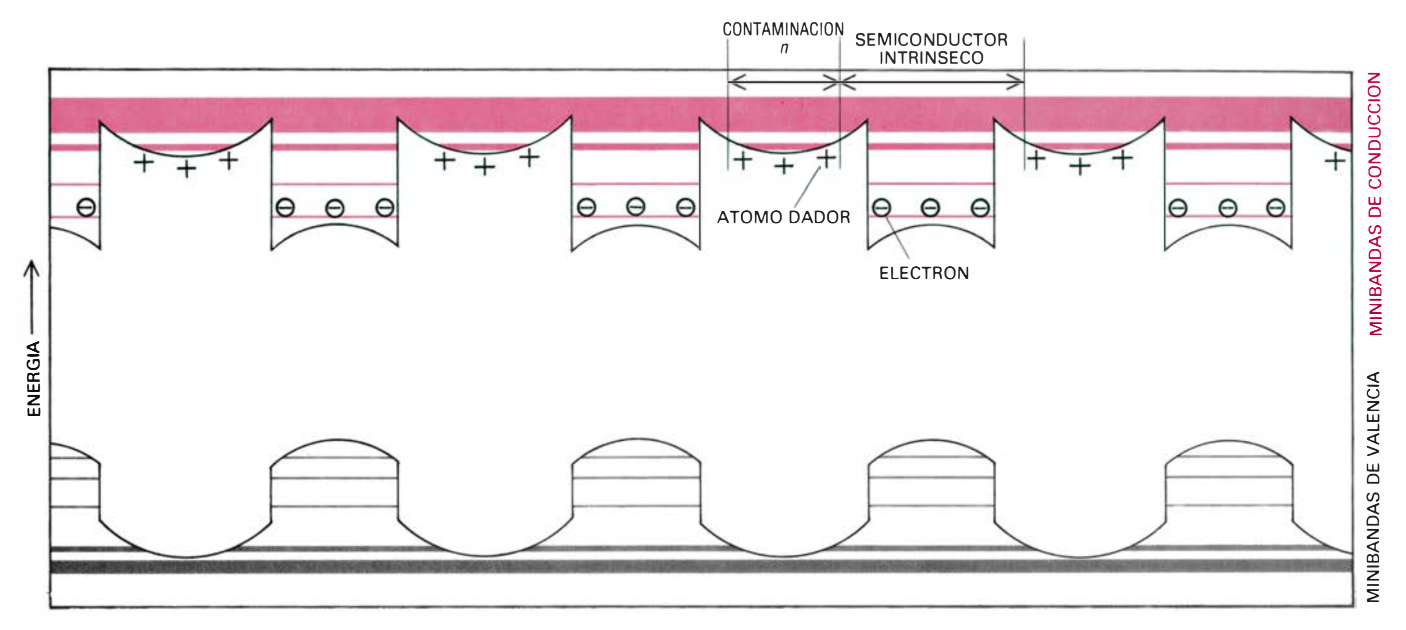
En este sentido, las superestructuras poseen también otras propiedades interesantes. Hasta aquí me he referido solamente a los materiales semiconductores en los que los portadores de cargas se crean por absorción de la luz. En lugar de dichos portadores, que desaparecen rápidamente debido a la recombinación entre los electrones y los agujeros cuando cesa la excitación del cristal, se puede introducir una población permanente de electrones o agujeros "contaminando" el semiconductor con átomos de impurezas. Los átomos







4. SE PRODUCEN LAS OSCILACIONES DE BLOCH cuando la aplicación de un voltaje acelera los electrones de la banda de conducción de un semiconductor; al mismo tiempo, producen una inclinación en dicha banda. Los electrones son entonces impulsados hacia el extremo superior de la banda. En un semiconductor típico (*a*) nunca alcanzarían ese lugar, sino que se produciría una emisión de fonones, es decir, los electrones originarían ciertas excitaciones térmicas en la red del cristal y "caerían" hasta una energía inferior. En cambio, en una superestructura, la banda de conducción se ha desdoblado en numerosas minibandas estrechas y sí que es posible dicha transición. Los electrones llegan al extremo superior de la minibanda de conducción más baja (*b*) y son reflejados, pudiendo oscilar repetidamente entre los dos extremos de esa minibanda antes de que se produzca la emisión de un fonón. Puesto que la minibanda está inclinada, dicha emisión desplaza el centro de masas del electrón cierta distancia Δz (*c*). Esta distancia disminuye al aumentar la inclinación de la minibanda. Por tanto, un aumento del voltaje puede producir una disminución de la intensidad de la corriente que atraviesa el cristal.



5. CONTAMINACION de una superestructura composicional (introducción de impurezas en la estructura cristalina); aumenta la movilidad de los electrones. El semiconductor en el que la separación entre las bandas de conducción y de valencia es mayor se contamina con átomos dadores, cada uno de los cuales contribuye con un electrón al cristal y, por tanto, se convierte en un ion positivo (*signos más*). Los electrones (*signos menos*) ocupan la minibanda de conducción más baja disponible; esta minibanda está en las otras capas. La separación espacial entre los portadores de carga y los iones positivos reduce su interacción; en virtud de ello, la conductividad del cristal centuplica varias veces la de un semiconductor uniforme que posea la misma concentración de portadores. La configuración que adquieren los pozos de potencial periódicos de una superestructura está producida por los electrones y los iones.

cuyo número de electrones de valencia (es decir, electrones que pueden formar parte de los enlaces químicos) supera en una unidad al número de electrones que poseen los átomos del cristal son dadores de electrones; estos átomos pueden ceder fácilmente el electrón en exceso menos fuertemente ligado. En consecuencia, un semiconductor cargado con átomos dadores ya no se comporta como un aislante, ni siquiera a temperaturas del cero absoluto; antes bien, se convierte en un conductor de electrones, o semiconductor de tipo *n*, en el que la letra *n* se refiere a la carga negativa de los electrones. Los átomos cuyo número de electrones de valencia es inferior en una unidad a los que poseen los átomos del cristal son aceptores de electrones; estos átomos aceptan fácilmente un electrón de valencia de dichos átomos del cristal y, en virtud de ello, se produce la formación de un agujero. De aquí que un semiconductor cargado con átomos aceptores sea un conductor de agujeros, o semiconductor de tipo *p*, en donde la letra *p* simboliza la carga positiva efectiva del agujero en cuestión.

Desgraciadamente, los átomos dadores y aceptores de un semiconductor no se limitan a proporcionar portadores de cargas. Al mismo tiempo, reducen la movilidad de dichos portadores. La explicación de este hecho radica en que cuando un dador cede un electrón, o cuando un aceptor lo gana, se convierte en un ion y crea a su alrededor, por tanto, un campo electrostático que impide el movimiento de los portadores de carga difundiéndolos. Para un ingeniero este hecho constituye un inconveniente, puesto que se traduce en que los electrones y los agujeros ya no se mueven tan deprisa como sería de desear. Para un investigador de los sistemas mecánico-cuánticos de muchos cuerpos se trata de otro dato negativo, pues la interacción entre los electrones resultará alterada por la interacción entre los electrones y los átomos de impurezas, cuya intensidad es comparable a la anterior.

En el caso de las superestructuras esta desventaja se salva mediante una estratagema muy simple, basada en la posibilidad de confinar la carga de los átomos dadores en aquellas capas del semiconductor donde la separación entre las bandas sea mayor [*véase la figura 5*]. Cada átomo dador aporta un electrón al cristal. Estos electrones tratarán de ocupar los estados vacíos cuya energía sea más baja. Sin embargo, dichos estados se alojarán en las capas de los otros semiconductores (allí estarán en las minibandas de conducción más bajas). En virtud de lo cual, la separación espacial de los electrones de los dadores será mayor y la conductividad del material en cuestión centuplicará varias veces la que tendría un semiconductor uniforme con la misma concentración de portadores. Esta estratagema fue comprobada experimentalmente, en 1978, por Horst L. Störmer, Dingle, Gossard, Wiegmann y Ralph A. Logan en los Laboratorios Bell. Una vez más, los ensayos confirmaron las predicciones. Las técnicas de crecimiento de cristales que se utilizan en la actualidad en los Estados Unidos, Europa y Japón proporcionan constantemente nuevos éxitos, que son bien recibidos por parte de los ingenieros interesados en el desarrollo de dispositivos electrónicos rápidos.

Cada vez que se produce un nuevo avance en las investigaciones relacionadas con las superestructuras de estado sólido crece mi convencimiento de que existe la posibilidad de crear otro tipo de superestructuras. Este nuevo tipo de material se podría obtener a partir de un semiconductor ordinario, con una modulación periódica de tipo *n* o *p*. El problema radica en despertar el interés de los investigadores capaces de preparar dicho material. En mi opinión, las dificultades técnicas que plantea la preparación de esa superestructura contaminada no serían mayores que las que presenta la obtención de las superestructuras ordinarias. En 1972 me ocupé ya de las propiedades electrónicas y ópticas de estas superestructuras, que no desmerecen de las que caracterizan a las superestructuras ordinarias.

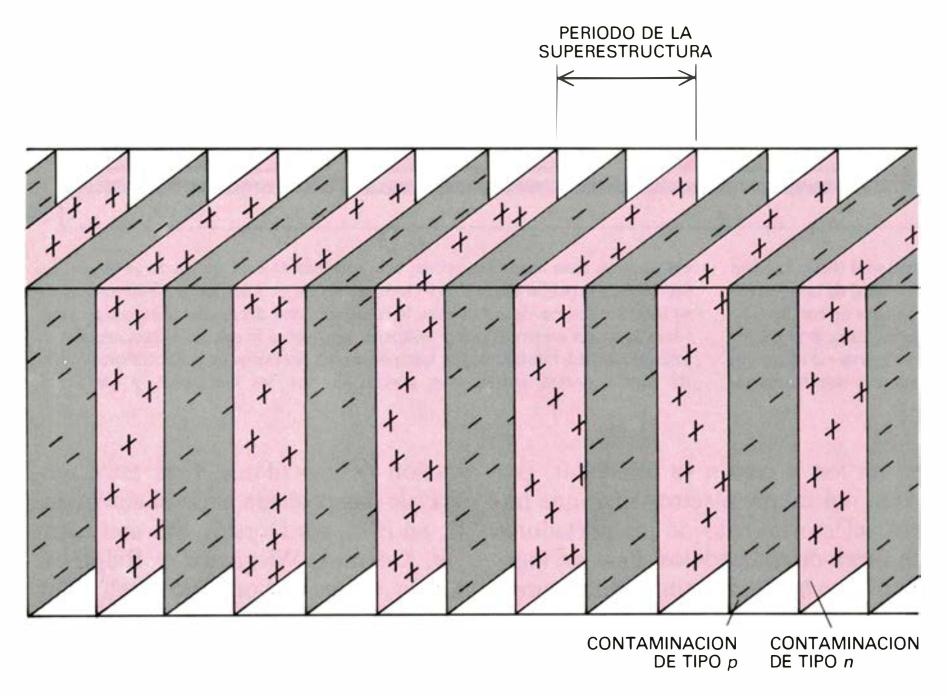
Vamos a considerar ahora un conjunto de capas de semiconductores

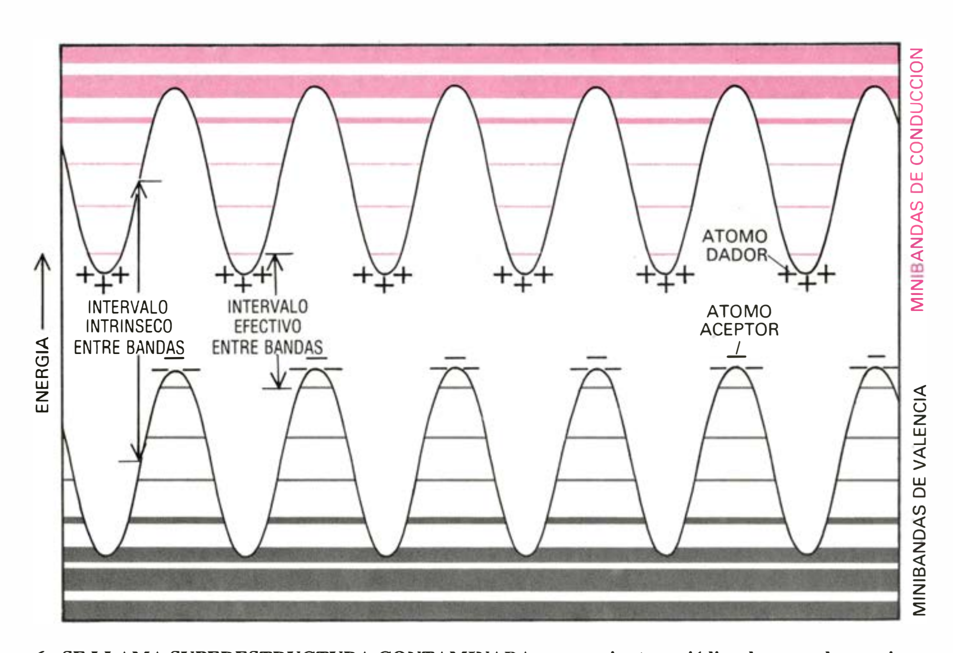
de tipo *n* y *p* separadas entre sí por unas capas no contaminadas del mismo semiconductor de base. (Estos últimos semiconductores son del tipo *i*, en donde la letra *i* representa la palabra intrínseco; la superestructura en cuestión será pues un cristal del tipo *n-i-p*.) Una cierta fracción de los átomos dadores y aceptores se ionizaría debido a que los electrones cedidos por los dadores se recombinarían con los agujeros que producen los aceptores en tanto el proceso tuviera lugar y disminuyera la energía total del material. La carga resultante de las capas contaminadas (con una carga positiva en cada capa *n* y una carga negativa en cada capa *p*) produciría un potencial electrostático periódico; este potencial modularía las bandas de conducción y de valencia de forma análoga a como lo hace en una superestructura de tipo composicional.

Una de las ventajas más importantes de las superestructuras contaminadas radica en la flexibilidad con que pueden modificarse sus propiedades. De hecho, sirve cualquier semiconductor de base para una superestructura contaminada, siempre que exista la posibilidad de obtener ambos tipos de cargas, la *n* y la *p*. En cambio, en el caso de una superestructura composicional, la elección de los dos semiconductores que la constituyen está fuertemente restringida por la exigencia de que las distancias interatómicas en uno de ellos se ajusten lo más posible a las distancias del otro. Eligiendo de forma adecuada la concentración de la carga y el espesor de la capa en cuestión, se puede lograr que la anchura de banda efectiva de una superestructura contaminada adquiera cualquier valor comprendido entre cero y la separación entre las bandas de valencia y de conducción que es característica del material original no contaminado. Si dicha concentración y espesor son suficientemente grandes, la superestructura se convierte en un semimetal, es decir, un material cuya anchura de banda vale cero y que posee determinada concentración residual de electrones libres y agujeros.

Existe la posibilidad de que las propiedades de una superestructura contaminada se modifiquen incluso después de haberla preparado. En particular, las propiedades eléctricas y ópticas de una superestructura contaminada se pueden variar dentro de límites muy amplios, sometiendo el cristal a una excitación débil; por ejemplo, mediante la absorción de una señal luminosa de baja intensidad o la aplicación de una corriente eléctrica poco intensa. Esta posibilidad, que distingue a las superestructuras contaminadas de todos los demás semiconductores, tiene su origen en la separación espacial entre los electrones y los agujeros que producen dichas excitaciones. Los electrones, situados ahora en las capas de tipo *n*, se dirigen hacia la minibanda de conducción más baja. Los agujeros, situados en cambio en las capas de tipo *p*, "ascienden" hasta la minibanda de valencia más alta.

Entre las capas *n* y *p* hay un hiato. En consecuencia, tanto los electrones como los agujeros carecen de las condiciones necesarias para que se produzca una recombinación. Si obedecieran las leyes de la física clásica, nunca se recombinarían, a menos que la energía térmica del cristal situara unos junto a otros. La recombinación por este mecanismo sólo es posible a temperaturas elevadas. Sin embargo, de acuerdo con



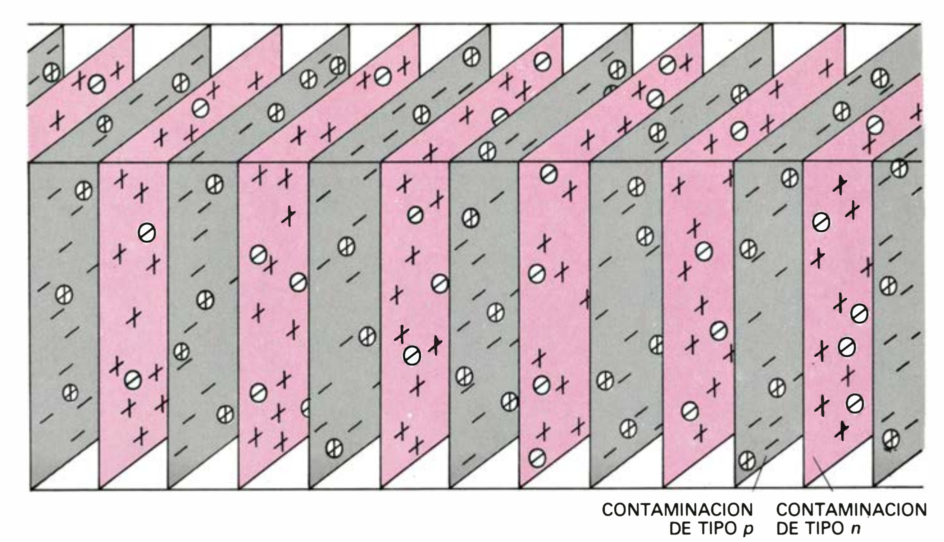


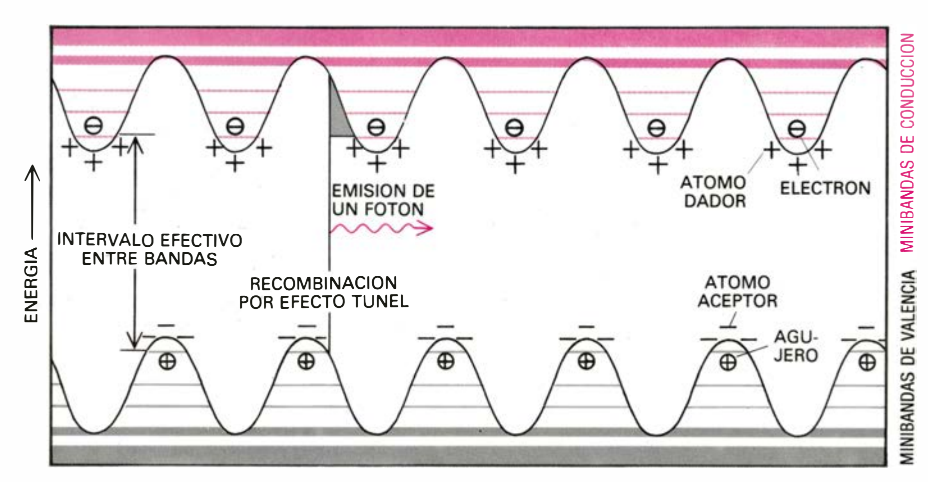
6. SE LLAMA SUPERESTRUCTURA CONTAMINADA a un conjunto periódico de capas de un mismo semiconductor contaminadas de dos formas diferentes (*dibujo superior*). En las capas *n*, los átomos dadores (*signos más*) aportan electrones. En las capas *p*, los átomos aceptores (*signos menos*) capturan electrones. La distribución de carga eléctrica resultante crea un conjunto de pozos de potencial (*esquema gráfico inferior*). El valor de la separación efectiva entre las bandas (diferencia entre las energías de la minibanda de valencia más alta y la minibanda de conducción más baja) depende de las concentraciones de los contaminantes, así como de los espesores de las capas que constituyen la superestructura en cuestión.

las leyes de la mecánica cuántica, existe otra forma de recombinación, basada en el principio según el cual las partículas pueden atravesar una barrera de potencial por "efecto túnel". En cualquier caso, la probabilidad de recombinación por efecto túnel disminuye espectacularmente al aumentar la altura y la anchura de la barrera. Los correspondientes cálculos demuestran que las superestructuras contaminadas se pueden diseñar de forma tal que las vidas medias de los electrones libres y los agujeros varíen entre algunos nanosegundos (es decir, un valor del orden de las vidas medias de los portadores de carga en los semiconductores uniformes) y varias horas.

De todo cuanto acabamos de exponer se desprende la conclusión siguiente: en una superestructura contaminada se puede conseguir que los estados situados muy por encima del estado fundamental (el estado de energía total más baja) se desintegren muy lentamente. De hecho, dichos estados se pueden estabilizar con sólo inyectar electrones y agujeros (la inyección de agujeros produce la extracción de electrones de la banda de valencia) o provocando su formación poco a poco. Por tanto, la concentración de electrones o agujeros en una superestructura contaminada es una magnitud susceptible de ser controlada con relativa facilidad. Este control es crucial. En las capas de tipo n, la concentración de electrones neutraliza en parte las cargas positivas de los átomos dadores ionizados. En las capas de tipo p, la concentración de agujeros neutraliza en parte las cargas negativas de los átomos aceptores ionizados. En virtud de ambos procesos, se reduce el potencial de la superestructura y aumenta la anchura de la banda efectiva.

Quizás el lector no se sorprenda demasiado si le decimos que la conductividad eléctrica, la absorción óptica, la emisión luminosa y, de hecho, prácticamente todas las propiedades eléctricas y ópticas de las superestructuras contaminadas pueden controlarse con gran facilidad. Los pares electrón-agujero creados por excitación óptica del cristal se consiguen separar entre sí con un alto grado de rendimiento. Dado que la recombinación es improbable, aumenta la concentración de pares electrón-agujero y, por tanto, el potencial se reduce y se amplía la anchura de la banda efectiva. La reducción del potencial incrementa la probabilidad de que se produzca el efecto túnel; por cuya razón, disminuye la vida media de los portadores de carga. Finalmente, llega un momento en que la velocidad a la que se produce la recombinación se iguala a la velocidad con que se crean nuevos pares electrón-agujero. Cuando se alcanza este equilibrio las propiedades del cristal se mantienen constantes e iguales a los valores deseados.

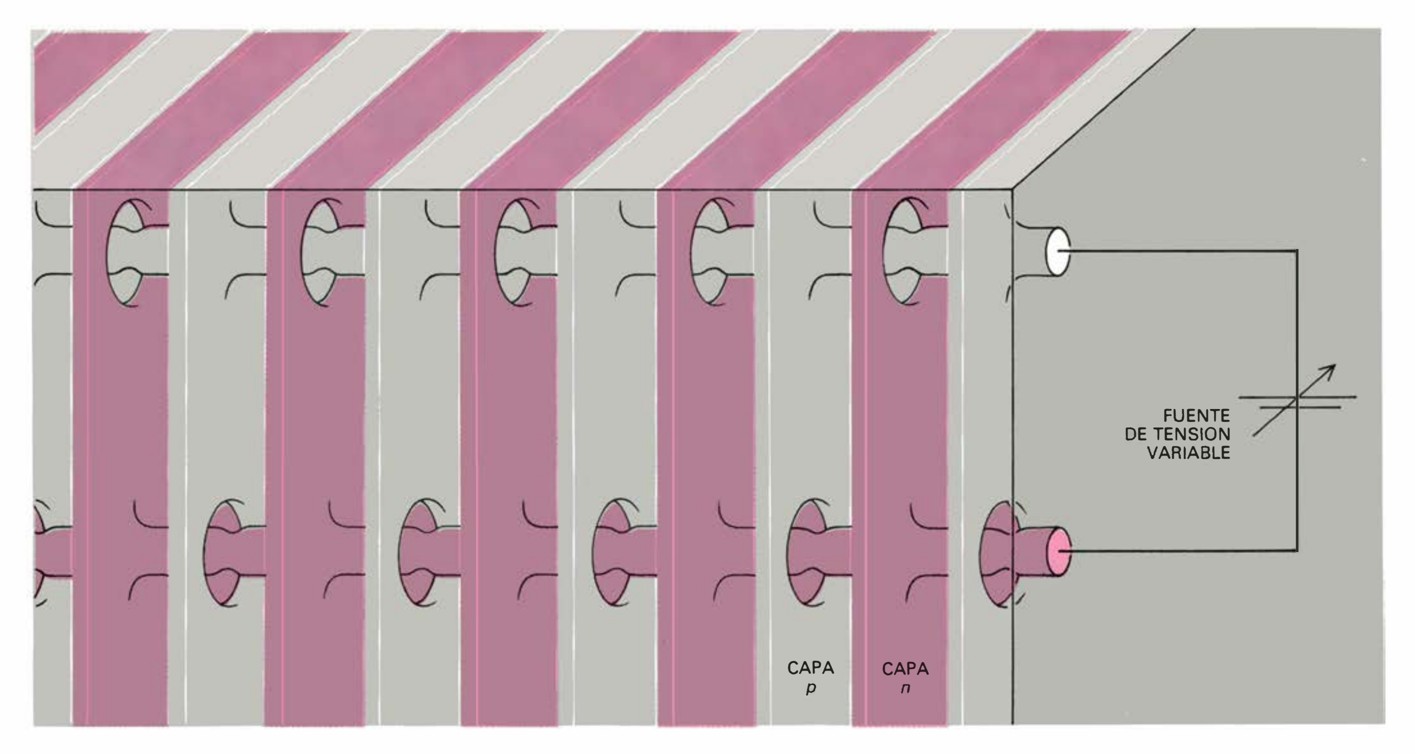




7. ESTADO EXCITADO de una superestructura contaminada ("doping"). Se produce cuando los electrones (*signo menos encerrado en círculos*) penetran en las capas *n* y los agujeros (*signo más encerrado en círculos*) penetran en las capas *p*. Estos portadores de carga neutralizan en parte la carga de los átomos contaminantes ionizados, de forma tal que los pozos de potencial se hacen más planos y aumenta la separación efectiva entre las bandas. Las propiedades de una superestructura contaminada pueden modificarse, pues, inyectando a su través una corriente eléctrica o excitando ópticamente los pares electrón-agujero. Los electrones y agujeros se pueden recombinar si los electrones consiguen atravesar la barrera de potencial (*en color gris*) por "efecto túnel". El rendimiento de este efecto depende de la anchura y la altura de la barrera; por tanto, la vida media de los portadores de carga también se puede modificar.

A primera vista, pudiera parecer que la excitación del cristal por inyección de electrones sólo en las capas n y de agujeros sólo en las capas p es un problema de microingeniería insoluble. La verdad es que la inyección no encierra mayor dificultad. Un electrodo selectivo para las capas n o para las capas p puede estar constituido por una región de fuerte contaminación de tipo n o p, respectivamente, extendiéndose a lo largo de toda la superestructura [*véase la figura 8*]. La introducción de portadores de carga en las capas contaminadas por cargas del signo opuesto se ve imposibilitada por la acción de los oportunos campos eléctricos. Concretamente, en las capas de tipo n los dadores ionizados repelen a los agujeros; en las capas de tipo p son los aceptores ionizados los que repelen a los electrones.

El investigador que finalmente emprendió la preparación de una superestructura contaminada fue Klaus Ploog, del Instituto Max Planck de Física del Estado Sólido de Stuttgart. Gracias a su habilidad, y a la de sus colegas, su trabajo se vio coronado por el éxito en el año 1980. El material base fue el arseniuro de galio; por contaminantes emplearon átomos de silicio, que actuaban como dadores, y átomos de berilio,



8. ELECTRONES SELECTIVOS CONECTADOS a las capas *n* y *p* de una superestructura contaminada; están constituidos por otras tantas columnas de un material fuertemente contaminado con impurezas de tipo *n* y *p*. Los dadores ionizados de las capas *n* repelen los agujeros del electrodo *p*; los aceptores ionizados de las capas *p* repelen los electrones del electrodo *n*. Como consecuencia, la aplicación de una diferencia de potencial entre dichos electrodos envía a los agujeros exclusivamente hacia las capas *p* y a los electrones exclusivamente hacia las capas *n*. Ello permite modificar fácilmente las propiedades del cristal. La inyección de portadores de carga cesa cuando la separación efectiva entre las bandas se iguala a la diferencia de potencial.

que lo hacían de aceptores. El método de crecimiento utilizado se fundaba en la técnica de los haces moleculares.

En primer lugar, se comprobó la relación predicha entre la concentración de los portadores de carga y la anchura efectiva de la separación entre las bandas midiendo la conductividad de los electrones y agujeros en las capas de la red cristalina. Un equipo de investigadores formado por H. Künzel, K. Ploog, J. Knecht, A. Fischer y el autor de este artículo encontró que la conductividad se anulaba para valores del potencial eléctrico para los que la concentración de los portadores era igual a cero. En segundo lugar, se comprobaron las propiedades ópticas predichas por la teoría. En una experiencia llevada a cabo por Künzel, Ploog y el autor, se demostró que la absorción de la luz por una estructura contaminada generaba portadores de carga con vidas medias superiores a los 1000 segundos. Esta cifra representa un aumento en las vidas medias de los portadores de carga en los semiconductores ordinarios en un factor del orden del billón. La sensibilidad de las superestructuras hacia la luz aumenta en un factor análogo. Evidentemente, la utilización de estas superestructuras contaminadas en la fabricación de detectores luminosos parece muy prometedora.

Estas investigaciones nos permitieron comprobar la posibilidad de hacer variar las propiedades ópticas de las estructuras contaminadas, tal como predice la correspondiente teoría. La prueba más contundente de dicha posibilidad la proporcionaron los experimentos realizados en relación con la luminiscencia de los cristales. En este sentido, los investigadores que he mencionado anteriormente, trabajando conjuntamente con H. Jung, D. Olego, E. Rehm y H. J. Stolz, midieron el espectro emitido por una superestructura contaminada bajo la excitación de diferentes intensidades luminosas. Al aumentar la intensidad de la luz incidente, la concentración de portadores, la anchura efectiva de la separación entre las bandas y, por lo tanto, la energía de los fotones emitidos por la superestructura, aumentaban hasta un 20 por ciento. Posteriormente se han observado aumentos similares a éste en el caso de superestructuras contaminadas excitadas por la inyección de portadores de cargas. Se plantea así la posibilidad de investigar la utilización de estas superestructuras como moduladores ópticos, es decir, en los dispositivos electrónicos que procesan haces láser.

Como es natural, la emisión de luz por parte de una superestructura exige la recombinación de electrones y agujeros, lo cual puede parecer incongruente con unas vidas medias de los portadores de carga del orden de 100 segundos. No hay tal incongruencia. En una superestructura contaminada se pueden dar dichas vidas medias cuando se encuentra en un estado fundamental y la anchura efectiva de la separación entre las bandas es prácticamente nula. (Con una separación entre bandas nula, la recombinación no puede disminuir la energía total del cristal; como consecuencia, las vidas medias de los portadores son elevadas.) Sin embargo, bajo condiciones de iluminación fuerte, la anchura efectiva de dicho intervalo se asemeja a la que suelen poseer los semiconductores no contaminados. La altura de la barrera interpuesta al efecto túnel disminuye y, por tanto, las vidas medias de los portadores de carga se hacen también menores.

Si se intenta producir una superestructura contaminada con luminiscencia elevada para energías de los fotones muy diferentes de las que corresponden al intervalo de bandas característico de los semiconductores no contaminados, se deberá proceder a un ajuste adecuado. Concretamente, se utilizarán capas delgadas y con una alta concentración de contaminante. Las capas delgadas suponen pozos de po-

tencial estrechos, lo que a su vez asegurará cierto grado de "transparencia" para los electrones libres del cristal incluso en el caso de que la altura de dichos potenciales sea notable.

En la lista de las propiedades de las superestructuras contaminadas predichas por la física cuántica y comprobadas experimentalmente con posterioridad se debe incluir, también, la verificación experimental de las diferencias en las energías de las minibandas. Los investigadores del Instituto Max Planck de Stuttgart, y en particular, G. Abstreiter y P. Ruden, hicieron incidir un haz de fotones sobre una superestructura contaminada de arseniuro de galio. Estos autores midieron la energía de los fotones difundidos por dicho material. Algunos de los fotones difundidos habían excitado a los electrones de las minibandas llenas a otras minibandas vacías más altas y, por consiguiente, habían perdido parte de su energía original.

El estudio de las superestructuras de estado sólido es uno de los campos de trabajo más importantes de la física del estado sólido en la actualidad. Los materiales semiconductores utilizados en la preparación de superestructuras composicionales ya no se limitan al arseniuro de galio y al arseniuro de galio y aluminio. Recientemente se ha preparado una superestructura composicional constituida por una serie de capas alternadas de arseniuro de indio (InAs) y antimoniuro de galio (GaSb); este material fue investigado originalmente por George A. Sai-Halasz, Chang, Tsu y Esaki en la empresa IBM. Las peculiaridades que presentan las propiedades electrónicas de esta superestructura arrancan de las posiciones, nada corrientes, de sus bandas. Sucede que el extremo inferior de la banda de conducción del arseniuro de indio está situado por debajo del extremo superior de la banda de valencia del antimoniuro de galio. Los portadores que coexisten en estos materiales mezclan sus características y se conocen con el nombre de hoctrones.

Las superestructuras de estado sólido son algo más que unos simples juguetes con los que se entretienen los físicos de estado sólido. En mi opinión, las superestructuras composicionales y, en particular, las superestructuras contaminadas, con sus posibilidades de ajuste sin igual, constituyen una nueva clase de semiconductores de enorme importancia, capaz de incidir en la física de estado sólido y en la electrónica.

Arqueología social de los monumentos megalíticos

El cambio producido en las estructuras funerarias del Neolítico de la Europa occidental, tendente a una mayor complejidad de las mismas, coincide con la aparición de un control político más centralizado

Colin Renfrew

A lo largo de los últimos veinte años muchos prehistoriadores han ido abandonando su pretensión de recrear el pasado en el marco de una historia de la cultura por un nuevo enfoque dinámico que no se limita a narrar los acontecimientos, sino que ahonda en sus causas y en su contexto. En palabras de Kent V. Flannery, de la Universidad de Michigan, "a algunos arqueólogos del Nuevo Mundo no les preocupa tanto el indio que hay detrás de un artefacto, producto de su actividad, cuanto el sistema que abarca al indio y el objeto. ¿Qué otros componentes tiene el sistema? ¿Qué recurso energético lo abastece...? Y demás cuestiones de ese jaez".

La diferencia entre ambos enfoques de la prehistoria se hace palpable en el Viejo Mundo a propósito de los megalitos europeos (del griego *megas*, "grande", y *lithos*, "piedra"). Estos impresionantes monumentos han intrigado a los estudiosos del pasado durante más de un siglo. Se encuentran en todos los países de la Europa atlántica, desde la costa mediterránea hasta Suecia; los huesos humanos que muchos de ellos contienen demuestran que se utilizaron como tumbas. Aunque los hay muy sencillos, no debió resultar fácil su construcción. Los dólmenes constan de tres o cuatro grandes bloques de piedra que soportan una gran losa, que hace las veces de cubierta. Persisten en Europa unos 50.000 dólmenes; la tarea de reunir los grandes bloques de piedra, así como el esfuerzo necesario para construir las rampas de tierra que permitieron colocar en su lugar exacto al bloque-techo, debieron ser prodigiosos. Otros monumentos son mucho más complejos que los dólmenes y su construcción sin duda requirió a veces una inversión de cientos de miles de horas-hombre. Como ejemplos se pueden citar los grandes sepulcros de corredor de Newgrange, en Irlanda, y Maes Howe, en las islas Orkney.

Ya los primeros investigadores de los megalitos europeos comprendieron que se trataba de monumentos anteriores a la civilización romana, e incluso prehistóricos. Sin embargo, la proporción y complejidad de algunas de esas estructuras, como las tumbas de corredor, hizo que los mismos investigadores se mostrasen reacios a admitir que estos monumentos megalíticos fueran producto de una labor exclusiva de las gentes bárbaras de la prehistoria europea. La ausencia de objetos metálicos indicaba que sus artífices habían sido sencillos agricultores de la Nueva Edad de Piedra; pero, antes del desarrollo de las técnicas de datación absoluta, su edad real sólo podía estimarse de forma aproximada.

¿Representaban esas construcciones la influencia de ideas más avanzadas, infiltradas desde áreas que mantenían culturas más desarrolladas, así Creta y Grecia, desplazadas al oeste hasta alcanzar España y desde allí al norte, a lo largo de la costa atlántica? Partiendo de este supuesto, el "difusionismo" consideraba que los bárbaros europeos se habían limitado a copiar, de las civilizaciones más precoces del Cercano Oriente, sus prácticas funerarias. Los adeptos a la escuela de la historia cultural llegaron a elaborar un mapa donde se recogían los sucesivos avances de las supuestas influencias mediterráneas.

El desarrollo del método de datación por carbono-14 acabó con la teoría difusionista. Pronto se demostró que los monumentos megalíticos de diversas zonas europeas precedían, en casi 2000 años, a muchos de sus supuestos antecesores del Mediterráneo. Alguno de los más antiguos, en la Bretaña francesa,se había levantado antes del 4500 a. C. No cabía duda de que tales monumentos, las más viejas estructuras de piedra que se conservan en el mundo, eran de origen europeo.

Los arqueólogos se enfrentaban a un doble problema: hallar una explicación de los orígenes de esos monumentos y ofrecer alguna hipótesis acerca de su función que justificara su existencia. Desechadas las distintas reconstrucciones históricas que se han ido ofreciendo ha de abordarse la cuestión bajo una nueva óptica; a saber: la que considera el tema megalítico como un proceso dinámico, un sistema de interacción entre los monumentos resultantes y las sociedades que los construyeron. Seguimos sin apenas saber por qué los megalitos surgen precisamente en ciertos lugares y no en cualquier otro. Algo se ha avanzado, empero, en los presupuestos sociales subyacentes.

Aunque los monumentos quizás hayan cumplido otras funciones, muchos de ellos se destinaron a sepulturas. En algunas zonas, sobre todo en Gran Bretaña, existen monumentos prehistóricos en piedra de otras clases, sobresaliendo los grandes círculos líticos de Stonehenge, Avebury y el anillo de Brogar, en las Orkney. En Carnac, Bretaña, se encuentran grandes "alineamientos" de piedras hincadas y de disposición paralela, que alcanzan una longitud de 1,2 kilómetros. El método de datación por carbono-14 y la variedad de monumentos existentes en distintas áreas abogan por cinco o seis lugares de origen, relativamente independientes unos de otros. Uno de esos lugares fue seguramente la península bretona. Otro Portugal y, posiblemente, España; un tercer grupo sería Dinamarca, probablemente Irlanda y quizás el sur de Inglaterra. Glyn Daniel, de la Universidad de Cambridge, ha sugeri-



1. ZANJA QUE RODEA EL ANILLO DE BROGAR, uno de los dos *henges* existentes en la isla central de las Orkney. La zanja se abrió de nuevo en el curso de estudios que el autor realizó en la localidad. Se demostró que correspondía a trabajos de cantería realizados sobre la roca madre de arenisca. La inversión de trabajo necesaria para cavar la fosa, del Neolítico tardío, de más de 300 metros en circunferencia, se calcula en 80.000 horas.

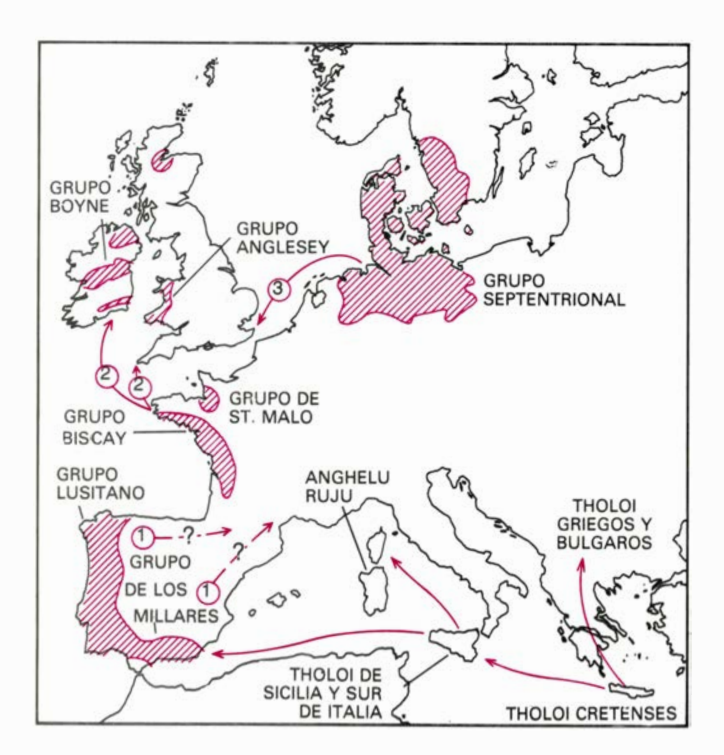
do que alguno de los monumentos tomara su forma por imitación en piedra de las casas de madera de los habitantes de esas zonas y de sus antepasados. Una idea plausible que no precisa, sin embargo, la razón de la existencia de esos monumentos.

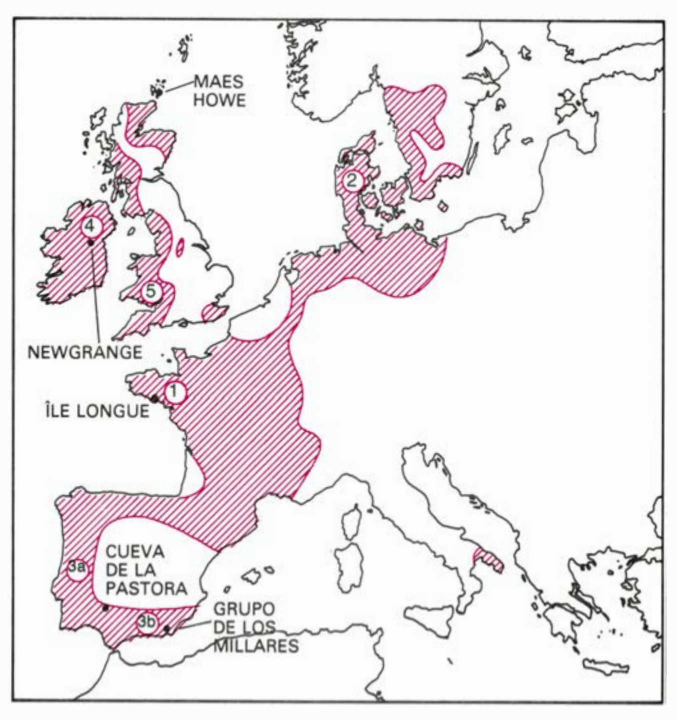
Ante un problema así, lo mejor que puede hacer el arqueólogo es inquirir por la función de los megalitos en la sociedad de su tiempo: ¿Cómo se utilizaban? ¿De qué modo facilitaban las labores de la sociedad? Y preguntarse por la misión es abordar también la naturaleza de la propia sociedad donde surgían. ¿Qué tipo de sociedades eran aquellas que concedían tan alto significado a estos monumentos en su seno? Es innecesario aclarar que ambas preguntas trascienden la idea "utilitarista" o funcional del megalito. En principio estas grandes proezas arquitectónicas debieron encerrar, en muchos casos, un enorme valor simbólico. Constituían seguramente un motivo de orgullo para sus constructores y quizá de envidia para sus vecinos. En un enfoque social de estos problemas hay que dar cabida al factor humano: orgullo, emulación, solidaridad y competición.

Al margen de su interés general, las dos preguntas deben responderse en la práctica y, en primer lugar, circunscritas a una región o lugar específicos. Los monumentos difieren de una zona a otra y deben interpretarse con tales especificidades, por más que sigan presentando rasgos comunes. En este sentido, una de las primeras áreas que se examinó fue la prehistórica región de Wessex, que se extiende por los condados del sur de Inglaterra, y donde se encuentran Stonehenge, Avebury y otros monumentos menores. El período se centra entre los primeros balbuceos de las comunidades agrícolas, allá por el 4000 a. de C., y el desarrollo de la tecnología basada en el trabajo del bronce, hacia el 2000 a. de C., cuando cesó la construcción de monumentos megalíticos. Hasta hace poco no se sabía gran cosa sobre los asentamientos domésticos en esa zona. Incluso hoy la mayoría de nuestros conocimientos proceden de la notoria profusión de monumentos, muchos de ellos asentados en las tierras cubiertas por margas, que fueron de las primeras en poblarse.

Los monumentos más antiguos de Wessex no son megalitos, es decir, no están formados por grandes piedras. Sin embargo, sí se trata de monumentos funerarios. Grandes montones de tierra de forma oblonga, y de más de 70 metros de largo, constituyen los llamados túmulos alargados (*long barrows*). Montones de tierras margas, son en algunos casos el equivalente local de las tumbas de piedra que comúnmente se encuentran en el oeste y, más escasamente, en la propia área de Wessex. Dado que los primeros investigadores no encontraron cámaras sepulcrales en ellos, se les denominó túmulos alargados sin cámara. Sin embargo, posteriores excavaciones revelaron que muchos de ellos habían tenido originariamente estructuras de madera, derrumbadas mucho tiempo antes. Al igual que las tumbas megalíticas con cámara, los túmulos cumplieron una misión sepulcral, aunque el número de individuos representado por los restos que nos han llegado es pequeño, en la mayoría de los casos. Además, los restos aparecen incompletos, lo que sugiere la hipótesis de que el difunto se depositaba en otro lugar para su descomposición, como preludio a su posterior enterramiento en el túmulo (una costumbre bien documentada en culturas de muchas partes del mundo).

Los objetos funerarios enterrados en los túmulos, muy elementales, son escasos, como ocurre en todas las tumbas megalíticas de Europa. Los hallazgos se limitan generalmente a los objetos propiamente neolíticos: cerámica, hachas de piedra pulimentada y herramientas de sílex tallado. Excepto en las tumbas de España y Portugal, pocas veces se





2. PRESUNTA DIFUSION de las tradiciones megalíticas en la Europa atlántica, tal y como se postulaba antes del empleo de los métodos de datación por carbono-14. Se basa tal hipótesis de la difusión en el supuesto de que los *tholoi* de Creta y Grecia, dotados de aleros, tejados cubiertos de tierra, serían anteriores a las tumbas similares de Italia y Sicilia, y que las más modernas habrían inspirado singularidades premegalíticas como los hipogeos de Malta y Cerdeña. Más tarde, la difusión de la tradición megalítica atravesaría la península Ibérica, hacia el oeste de Francia, norte de Alemania, Escandinavia y, finalmente, llegaría a Inglaterra, Gales, Escocia e Irlanda (*mapa de la izquierda*). Con el desarrollo de la datación absoluta, que demostró que ciertos monumentos megalíticos de Francia se habían construido hacia el 4500 a. de C., y que incluso los monumentos existentes en las remotas islas Orkney eran varios siglos más antiguos que las pirámides de Egipto, la interpretación difusionista dio paso a la hipótesis de los orígenes locales (*mapa de la derecha*).

entierran objetos ricamente labrados, como esculturas de piedra, piedras decoradas con incisiones o motivos de otro tipo. Puede, por tanto, aventurarse, para Wessex y para un área mayor, que las tumbas no servían de ostentación de la riqueza personal del difunto y que los objetos incluidos en ellas no son indicativos de un rango social alto. Ello contrasta notablemente con los hallazgos procedentes de la primera Edad del Bronce de Wessex. En los siglos inmediatamente posteriores al 2000 a. de C., las tumbas se acompañaban de ricos objetos funerarios: objetos de bronce, ámbar y oro.

Una segunda clase de monumentos de Wessex son recintos delimitados por zanjas concéntricas y discontinuas, conocidos entre los arqueólogos por *causewayed camps*. Muchos de ellos se encuentran burdamente vallados por una muralla circular de unos 200 metros de diámetro, inscrita en los círculos de fosos concéntricos. Las porciones no excavadas de las zanjas servían de zonas de paso. Sabemos ya que el elemento principal del monumento lo constituían los bancales construidos con yeso extraído de las zanjas, y no los propios fosos. Recientes excavaciones efectuadas en Hambledon Hill, Dorsetshire, por Roger Mercer, de la Universidad de Edimburgo, han demostrado que el recinto estaba defendido por una empalizada próxima a él. Mercer encontró también pruebas (cráneos humanos y fragmentos de esqueletos) que permitieron ampliar las primeras interpretaciones de esos asentamientos como lugares de encuentro de las comunidades dispersas por las zonas circundantes. Mercer sugiere que los recintos servían de "centro de exposición de cadáveres... un vasto y maloliente cementerio al aire libre cuyo silencio sólo rompían los graznidos de los cuervos y aves de rapiña".

Tanto los túmulos alargados como los círculos de zanjas corresponden a la primera parte del Neolítico británico, desde aproximadamente el 4000 a. de C. hasta el 3500. Más tarde, hacia el 3000 a. de C., hizo su aparición otro tipo de monumento: el *henge*, término derivado del famoso Stonehenge de Wessex. Muchos *henges* no son, sin embargo, círculos de grandes piedras, sino recintos circulares rodeados por una fosa con un terraplén bien asentado, dispuesto generalmente en el exterior de la fosa. Con una o dos vías de acceso, se distingue por un ordenamiento mucho más regular que el de los primeros círculos de zanjas. Las excavaciones llevadas a cabo en varios *henges* (en particular las de Durrington Walls, acometidas por Geoffrey Wainwright, inspector de Monumentos Antiguos del Departamento del Medio Ambiente del Reino Unido) han demostrado que los había con grandes cabañas circulares de madera. A veces, los *henges* son de grandes dimensiones, hasta de 500 metros de diámetro.



3. SILBURY HILL, un gran túmulo cercano a Avebury, en el condado de Wiltshire. La tumba, que ocupa el centro de esta fotografía aérea, fue, hasta la revolución industrial, la mayor obra humana de Europa; requirió una inversión aproximada de 18 millones de horas de trabajo humano para levantarla.

No pueden olvidarse en esta breve revisión de los monumentos de Wessex al propio Stonehenge, con su círculo de megalitos enlazados por dinteles, ni al colosal túmulo artificial de tierra de Silbury Hill, próximo a un círculo de piedras mucho mayor que Stonehenge: Avebury. Hasta la revolución industrial, Silbury Hill constituyó la mayor estructura artificial europea.

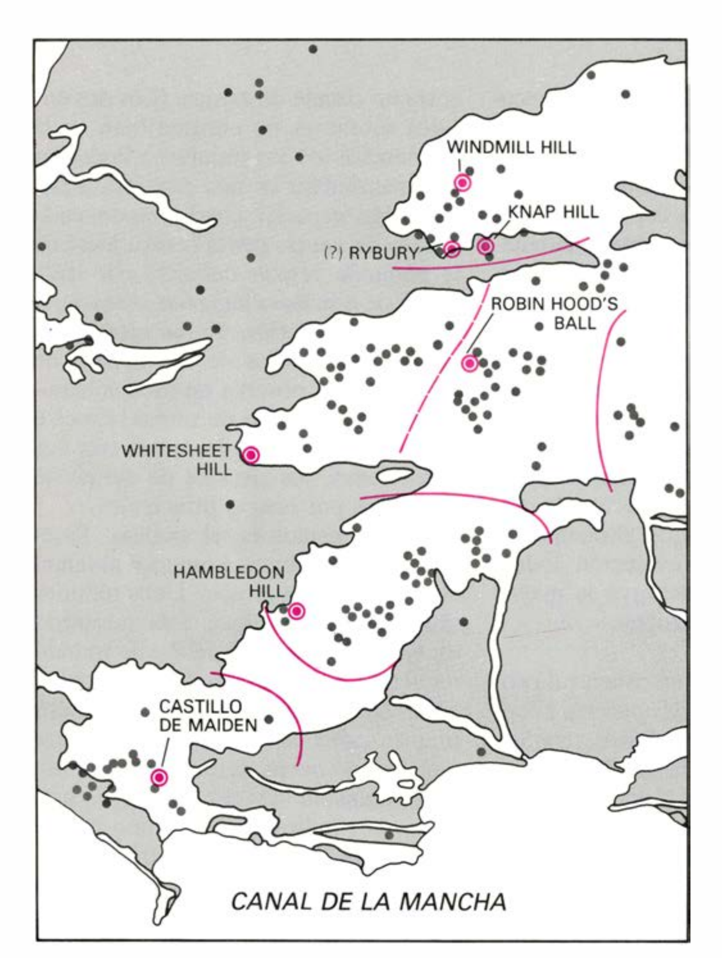
A la hora de poner en orden tal profusión de materiales pueden adoptarse dos enfoques analíticos: confeccionar un registro de la distribución espacial o considerar los términos absolutos. Ambos métodos mostrarían elementos jerarquizados, lo que, a su vez, reflejaría la jerarquización social subyacente dentro de la propia sociedad. Por ejemplo, al cartografiar los monumentos de Wessex se aprecia que los 132 túmulos alargados superan en número a los 7 círculos de zanjas (*causewayed camps*) en una proporción cercana a 20 por uno. El mapa muestra también una gran dispersión de los túmulos alargados, que pueden agruparse en cinco conjuntos. En cada conjunto entra un círculo de zanjas. (Los dos círculos sobrantes no encajan bien en la distribución de los túmulos.) Podemos entonces hablar de una modesta jerarquización espacial considerando cada círculo de zanjas como centro local de la pequeña región definida por cada grupo de túmulos alargados. Atendiendo a la distribución de los principales *henges* (los de más de 200 metros de diámetro), se observa un modelo similar al de los círculos de zanjas. En este período posterior, y de nuevo con dos excepciones, los círculos de zanjas se sustituyen por *henges* principales.

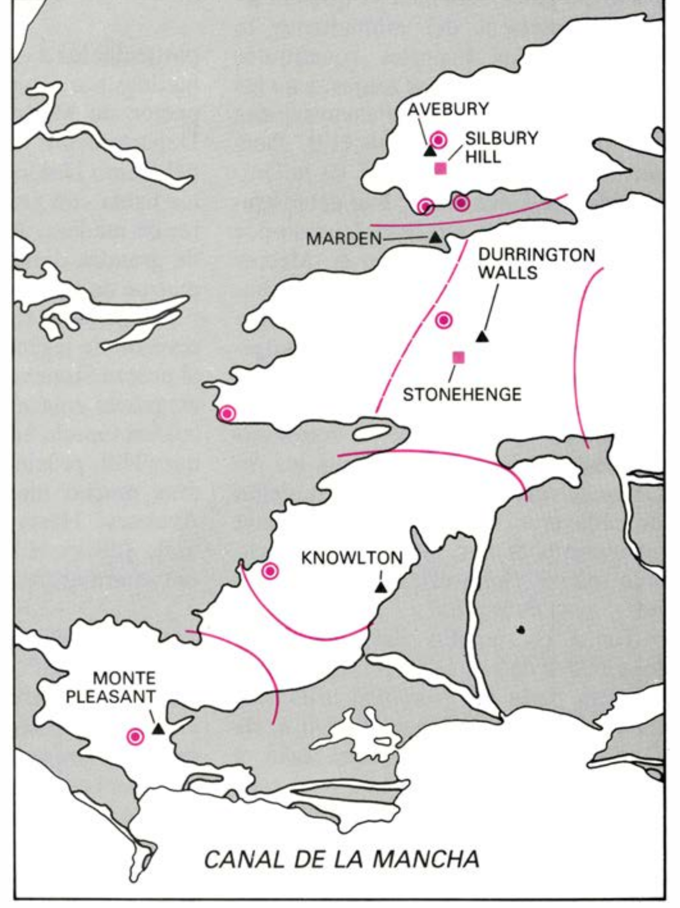
Profundicemos en el análisis. Pasemos de la distribución espacial al tamaño. ¿Qué se nos ofrece? Unos túmulos alargados de tamaño más bien modesto. El cálculo de la inversión de trabajo requerida para su construcción (resultante de estimaciones basadas en los simples métodos de remoción de tierra y el uso de herramientas elementales, como palas de madera y picos de asta de ciervo) indica que cada uno habría necesitado de 5000 a 10.000 horas-hombre de trabajo. Ello supone una inversión de trabajo comprendida entre las 10^3 y 10^4 horas-hombre. En otros términos, la cantidad de trabajo que 20 hombres podrían hacer en 50 días. El trabajo necesario para construir los *círculos de zanjas*, siguientes en el orden jerárquico, alcanza mayores proporciones. Podrían haberse necesitado de 40.000 a 100.000 horas-hombre, con lo que se alcanzan las 10^5 horas-hombre.

Los grandes *henges* del Neolítico tardío representan de nuevo un aumento de un orden de magnitud en la inversión de trabajo, aproximándose a las 10^6. En este sentido, Wainwright ha calculado que se dedicaron un total de 900.000 horas-hombre a la construcción de Durrington Walls. Y si los grandes *henges* constituyen las "figuras estelares" de los monumentos prehistóricos, Wessex poseería dos monumentos de primer rango: Silbury Hill y Stonehenge. Richard J. C. Atkinson, del University College de Cardiff, ha estimado el trabajo necesario para la construcción de Silbury Hill y Stonehenge (incluidas las tareas de transporte de la piedra hasta Stonehenge) en 18 y 30 millones de horas-hombre, respectivamente. Ello excede, con mucho, la magnitud de 10^7.

Existe, pues, una jerarquía de horas de trabajo humano junto al evidente escalonamiento espacial. Hace algunos años, sugerí que podíamos postular unas estructuras sociales relacionando estas jerarquías de espacio-energía, lo que, para el Neolítico tardío, nos autorizaría a hablar de sociedades jerarquizadas y centradas en los cinco mayores *henges* de Wessex. Los territorios que le corresponderían a cada uno de los *henges* puede inferirse, de una manera aproximada, dividiendo el terreno de suerte que cada *henge* principal constituya el foco dominante de su contorno. Señalé que para conseguir tan impresionante inversión de trabajo había que suponer la existencia de alguna autoridad que fuera el centro orgánico del poder, y que podían enumerarse los rasgos que caracterizan a las sociedades de estructura jerárquica de acuerdo con los parámetros de que hablan los etnólogos. Entre estos rasgos destaca la propia existencia de una escala social (con un jefe a su cabeza) que se asienta en un sistema basado en la redistribución, organizada por el jefe, de lo producido, incluidos bienes necesarios para la subsistencia. Las sociedades jerarquizadas se distinguen por una alta densidad de población y unos límites territoriales definidos de un modo más tajante que los de las sociedades paritarias. Poseen centros que coordinan las actividades sociales, religiosas y económicas; se caracterizan por un mayor desarrollo de la especialización; propio de ellos es también la capacidad para organizar y desplegar esfuerzos que sirvan a la consecución de empresas de gran alcance.

Alguno de estos rasgos podrían aplicarse a comunidades asentadas en centros como Durrington Walls o Avebury. El paso siguiente hacia una mayor centralización pudo traer la amalgama o asociación de territorios antes independientes formando una federación, quizá con Stonehenge como su centro religioso y Silbury Hill como su mayor monumento. Atkinson, apoyado en los datos arrancados de sus





4. MONUMENTOS FUNERARIOS del antiguo Wessex (los actuales Dorsetshire y Wiltshire), Inglaterra. Se agrupan en las desforestadas campiñas cubiertas de margas (*áreas blacas*) preferidas por los agricultores neolíticos que primero ocuparon la región. Sus túmulos funerarios (*puntos*) eran grandes montones de tierra; no constituían, por tanto, monumentos megalíticos propiamente dichos. Al agruparlos en cinco o seis conjuntos (*líneas de color*), cada agrupación recoge un solo círculo de zanjas (*causewayed camp*) (*puntos de color*), que, según parece, era un recinto donde se exponía a los difuntos antes de enterrar definitivamente sus restos en el grupo de túmulos alargados (*long barrows*) (*mapa de la izquierda*). Los objetos funerarios encontrados en la excavación de los túmulos alargados son escasos y sencillos, lo que sugiere que los agricultores formaban grupos sociales igualitarios. Cuando, medio milenio después, se construyeron en la misma zona de Inglaterra nuevos tipos de *henges*, que probablemente presentaran estructuras de madera dentro de sus límites, varias de las estructuras mayores (*triángulos del mapa de la derecha*) aparecen en los territorios definidos por los túmulos alargados, como había ocurrido antes con los círculos de zanjas. Ello aboga por una organización social que había ido escalonándose hacia una forma jerárquica. Otra prueba en este sentido la tenemos en la aparición, en esa zona, de dos monumentos "estelares": Stonehenge y, al norte, Silbury Hill (*cuadrados de color*).

excavaciones en Silbury Hill, sostiene que el monumento se construyó en sólo dos años. Aun contando que los obreros trabajaran ocho horas diarias durante 300 días al año, se habría necesitado un equipo de 3700 hombres a jornada completa para levantar el gran túmulo; ejército al que habría tenido que alimentar la organización de su respectivo grupo.

Este enfoque, que podría denominarse de "arqueología social", requiere la interpretación de los grandes monumentos dentro de un contexto social bien definido. Los sitúa dentro de un proceso dinámico en el que las sociedades más simples del período primario se transforman, en épocas posteriores, en unidades sociales más complejas, que pudieron tener una mayor densidad de población y, por consiguiente, mayor contingente humano. Permite también este enfoque separar el tema de los grandes monumentos (los grandes *henges* y Silbury Hill), que podemos considerar producto de las sociedades jerarquizadas, de aquellos monumentos menores y más locales, como son los túmulos alargados. El hecho reviste un obvio interés, ya que la mayoría de los megalitos europeos son propiamente equiparables a los túmulos alargados. Se trata también de monumentos funerarios, que requerían muchas veces casi 10.000 horas-hombre de trabajo para su construcción, y cuya dispersión imita la de aquellos túmulos. Los monumentos mayores son más escasos: es difícil encontrar *henges* y círculos de piedras en Europa continental (aunque, sin duda alguna, el gran alineamiento de Carnac pertenece a la categoría de Stonehenge).

De nuevo nos resultan útiles los conceptos de escala y organización espacial al considerar más detenidamente los túmulos alargados de Wessex, u otros monumentos de piedra de la misma época, como las tumbas megalíticas de la isla Arran, frente a la costa occidental de Escocia. En ambos casos, si se trazan líneas para dividir el terreno entre los monumentos y separarlo en unidades (proceso que los geógrafos llaman "formar polígonos Thiessen"), se observa, con raras excepciones, que el terreno se parcela en unidades productivas de tamaño similar. En Arran, cada polígono contiene por lo menos algún bancal de tierra arable. Dejando aparte los casos en los que los monumentos se encuentran a pares, el modelo también es disperso. Si a ello unimos el tamaño de los monumentos, no se encuentran indicios de estructura jerarquizada alguna: al menos en un orden de magnitud, las tumbas son de un tamaño aproximado. El modelo no es jerarquizado, sino paritario.

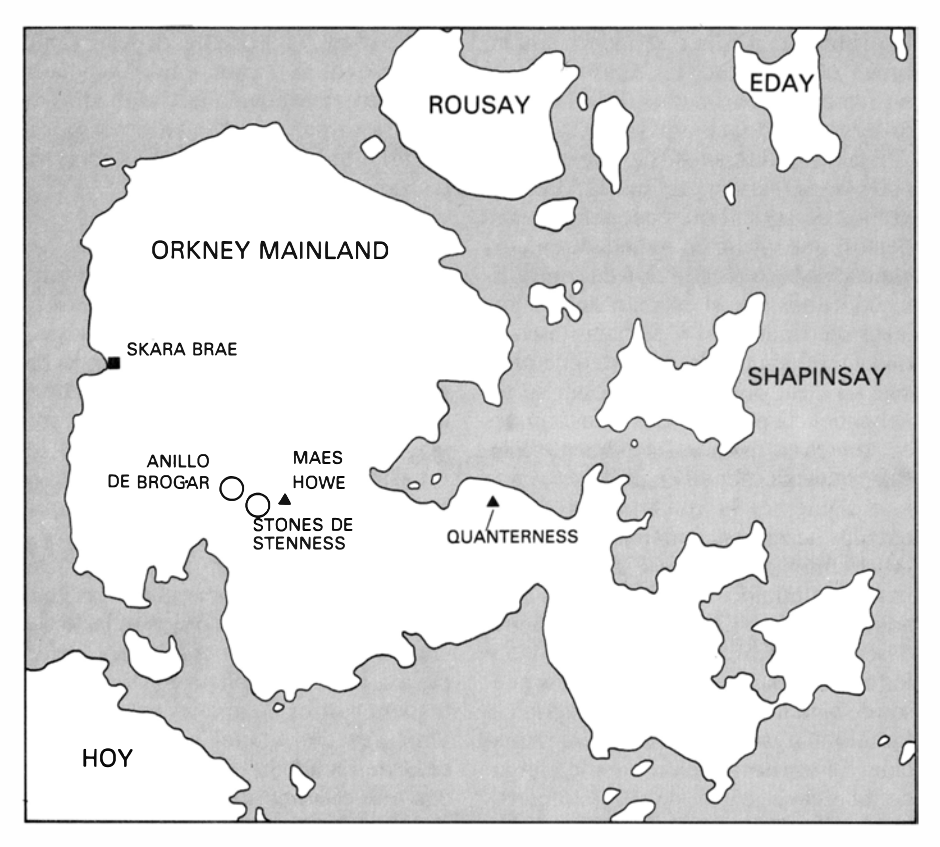
Estos monumentos, de hace 5000 o 6000 años, sirven de indicadores de grupos de agricultores pequeños y dispersos, que vivían en sociedades no organizadas bajo ningún tipo de centralismo. Grupos que sí estaban unidos por lazos matrimoniales y de parentesco (y quizá también por intercambio de mercancías), sin por ello abandonar su independencia política. Los lazos tribales existentes entre ellos, inicialmente a escala reducida, tienen un indicador en la fase siguiente: la construcción de los círculos de zanjas. Podemos imaginar a los difuntos de los más o menos 20 territorios definidos por los túmulos alargados de Hambledon Hill, vivamente descritos por Mercer, transportados a los círculos para su exposición; los huesos de algunos de ellos se devolverían al túmulo alargado local para su inhumación. El territorio de un túmulo alargado abarcaría cerca de 10 kilómetros cuadrados, tomando una cifra modesta de entre 10 y 50 hectáreas de tierra para mantener a cada persona en una economía agrícola rudimentaria; cabe entonces suponer, en términos demográficos, una población de entre 20 y 100 personas por territorio. Estas pequeñas unidades pueden considerarse "sociedades segmentarias", empleando la expresión de la antropología social; cada segmento sería un grupo de personas, económica y políticamente autónomo, que se autoabastecería y ejercería un control eficaz sobre sus recursos productivos.

Desde esta perspectiva, se aprecia más claramente el significado de las tumbas de cámara, mucho más ambiciosas. Se trataba, por supuesto, de lugares de enterramiento. De hecho, con la posible excepción de los túmulos alargados con cámara rodeados de una empalizada, fueron lugares de inhumación colectiva. Sin embargo, un grupo de quizá 50 personas, aunque ocupara de forma permanente durante siglos la misma porción de tierra, no necesitaría, para enterrar a sus difuntos, una tumba tan monumental como esas. Nada más sencillo que enterrarlos rápidamente sin más florituras. Por consiguiente, la monumentalidad de las tumbas con cámara no se limita lógicamente a su función funeraria. Induce a pensar en otros fines. En palabras de Andrew Fleming, de la Universidad de Sheffield, se trataba de "tumbas para la vida", lo que viene a decir que su carácter monumental es resultado de actividades deliberadas de los miembros de la sociedad, y que debieron servir a los objetivos de la misma. Para dilucidar el papel que desempeñaban las tumbas de cámara se han formulado tres explicaciones diferentes, que no son excluyentes, sino complementarias.

Avancé yo la primera explicación del porqué de esos ambiciosos monumentos; veía la razón en el ámbito de lo que podríamos llamar cohesión social. En épocas de tensiones, particularmente en circunstancias de alta presión demográfica, la comunidad actuaría con mayor resolución cuando todos sintieran conciencia de sus lazos sociales. La construcción de un monumento perdurable sería el mayor acto simbólico que afirmara con fuerza la unidad del grupo. Debe recordarse que estas gentes vivían en un mundo donde las creaciones humanas perdurables eran pocas: las casas solían construirse de madera y otros materiales perecederos. Una vez levantado el monumento, constituiría además el hito más elevado con que el hombre definiría su territorio. Incluso existiría el puntillo entre los vecinos por edificar el centro más grande y eminente.

Por ello, sostengo que deberíamos contemplar estos monumentos no como simples tumbas, sino como centros públicos, donde se reuniesen, además, en muchas ocasiones: quizás el lugar idóneo donde practicar ritos religiosos que pusieran en comunicación a la comunidad con sus antepasados y con sus más recientes difuntos. La construcción de un monumento focal podría ser la expresión visible de las aspiraciones del grupo local, sirviendo para garantizar la continuidad y cohesión del mismo, y, por consiguiente, su supervivencia.

La segunda explicación se halla en esta misma línea de razonamiento. Su autor, Arthur A. Saxe, de la Universidad de Ohio, pondera la relevancia de tales prácticas funerarias para el mantenimiento de las tierras ancestrales de la comunidad. "Hasta cierto punto, escribe, este colectivo social tiene derecho a usar o ejercer un control básico sobre los recursos limitados obtenidos o legitimados por medio de la descendencia lineal del difunto (es decir, lazos lineales con los antepasados); tales grupos mantendrán áreas destinadas exclusivamente a sus muertos." Así, el título de tierra comunal queda atestiguado por el mantenimiento de la tumba ancestral. Ambos enfoques pueden resumirse diciendo que los monumentos servían de marcas de delimitación del territorio de las sociedades segmentarias.



5. CUATRO MONUMENTOS MEGALITICOS, dos *henges* y dos sepulcros de cámara, se cuentan entre los ocho mayores de las islas Orkney. El autor y sus colegas reabrieron la tumba de cámara Quanterness, construida hacia el 3200 a. de C.; había sido parcialmente excavada hace dos siglos. También desenterraron el perímetro del anillo de Brogar, un monumental *henge* construido hacia el 3000 a. de C., quizá dos siglos antes de que se levantase el sepulcro de cámara de Maes Howe, más complejo. La sencillez de Quanterness indica que los agricultores que la levantaron constituían una pequeña comunidad.

La tercera explicación, postulada por Christopher Tilley, de la Universidad de Lund, ofrece un enfoque neomarxista que asigna una función más a los monumentos, función que no contradice las dos ideas precedentes. Tilley abunda en el papel del rito mortuorio, común a otros ciclos rituales de la vida, como "legitimación de intereses parciales" dentro de la sociedad. Dentro del propio marco más o menos paritario de las sociedades segmentarias, Tilley señala que, en las sociedades organizadas por linajes, hay clases contrapuestas, contraposición que se manifiesta sobre todo entre los más viejos y los miembros más jóvenes del grupo. Arguye que el control sobre el ritual, destacado por los monumentos, ayudaba a mantener el continuismo del dominio de los mayores. No estoy seguro de que tal argumento aporte mucho más que el punto de vista que considera los monumentos símbolos de cohesión social, aunque Tilley base su idea en un marco de razonamiento diferente.

Las tres explicaciones son funcionalistas, lo que equivale a decir que ninguna de ellas explica efectivamente el origen del megalitismo. Demuestran, sin embargo, que los monumentos desempeñaban una función útil y significativa en las sociedades que los erigieron. Este aspecto del significado de los megalitos debe recogerse en cualquier explicación que se pretenda satisfactoria, y que, ciertamente, debe ser una explicación ofrecida desde presupuestos sociales.

Nadie esconde cuán difícil resultaría comprobar directamente esas ideas con el objeto de poder establecer hipótesis y generalizaciones del estilo de las que nos ofrecen las ciencias exactas. Pero pueden contrastarse con los datos que aportan los nuevos trabajos de campo. Con esta intención emprendí, entre 1972 y 1974, unas excavaciones en las islas Orkney, las más septentrionales de las islas británicas.

Las Orkney se encuentran desforestadas desde la última glaciación, debido a los fuertes vientos que las azotan. Abunda la roca arenisca, de fácil fracturación laminar. La población prehistórica contaba, pues, con un abundante y útil recurso para la obtención de materiales de construcción al margen de la madera. Todas las construcciones de las Orkney se hicieron de arenisca, lo que explica su buen estado de conservación. Por otro lado, las prácticas agrícolas en las islas no se mecanizaron hasta muy recientemente, lo que ha impedido la demolición de ciertos túmulos de piedras, conocidos entre los especialistas por el término anglosajón de *cairns*. Tales túmulos son hoy claros indicadores de restos de enterramientos prehistóricos.

La fama de la antigüedad prehistórica de las islas Orkney arranca de lejos, en particular la relativa a sus dos *henges* con un círculo de piedras clavadas alrededor: el asentamiento de Skara Brae y el espléndido sepulcro de corredor de Maes Howe. Skara Brae, excavada en la década de 1930 por V. Gordon Childe, es una aldea magníficamente conservada del tiempo de los primeros agricultores de la isla (hacia el 3000 a. de C.). Las paredes de las casas conservan todavía una altura de más de dos metros; en algunas habitaciones persisten camas y alacenas de piedra. Maes Howe tiene una impresionante cámara construida en piedra a la que se accede por un corredor. Las piedras, cuidadosamente talladas, revelan un trabajo de óptima calidad; resulta difícil creer que no se utilizaran herramientas metálicas para su elaboración. En la vieja teoría difusionista sobre el origen del megalitismo, se consideraba Maes Howe el monumento más antiguo, por ser el monumento más perfecto de las Orkney. Se trataría de la primera factura de los supuestos nuevos emigrantes que desembarcaron aquí con sus ideas de procedencia, antes de su prolongada estancia en el remoto norte.

Cuando inicié mi trabajo, no se disponía de datos por carbono-14 para las Orkney. Me propuse, en consecuencia, obtener muestras de materia orgánica fechable, a fin de revelar la cronología de los sepulcros de cámara de las Orkney. El principal yacimiento elegido para excavar fue la tumba de Quanterness. Gracias a un informe escrito en el siglo XIX tuve noticia de que se había abierto Quanterness, para cerrarlo después. Mis colegas y yo esperábamos encontrar pruebas para la datación de su construcción y avanzar en los estudios con el fin de contribuir a la comprensión general de los monumentos de las islas. Coincidieron por las mismas fechas otras campañas de excavación en Skara Brae, dirigidas por David Clarke, del Museo Nacional de Antigüedades de Escocia. A su vez, Graham Ritchie, de la Real Comisión para los Monumentos Antiguos e Históricos de Escocia, empezaba a trabajar en uno de los *henges* de las islas, el Stones de Stenness.

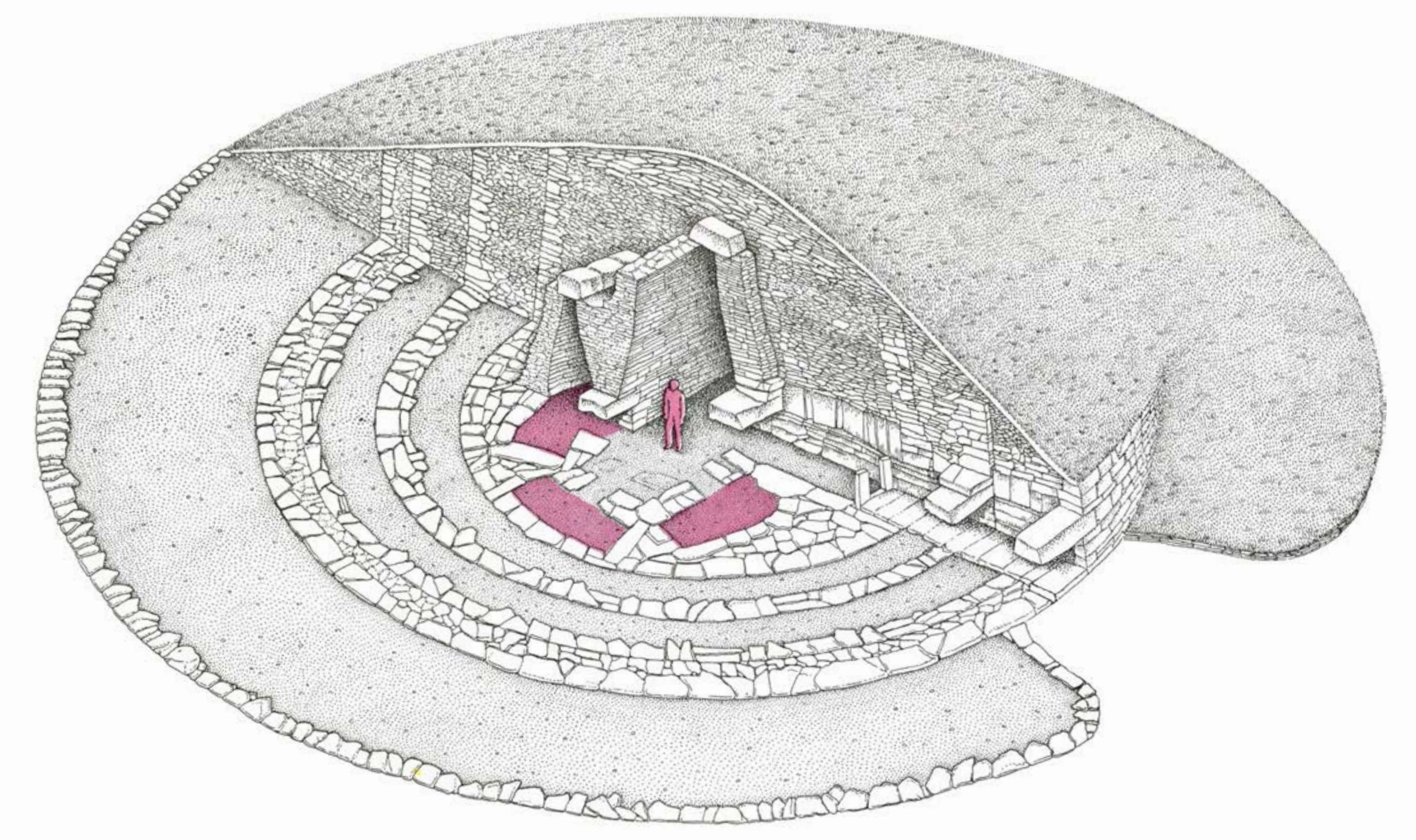
La techumbre de Quanterness se había derrumbado y hubimos de entrar a la cámara por arriba. Tal y como indicaba el croquis del siglo XIX, había seis cámaras guarneciendo la cámara rectangular principal, y un pasadizo de entrada a la misma por el este, ahora bloqueado. Sin embargo, aquel primer relato no hablaba del excelente trabajo de la piedra, labor que erige a este monumento en la más impresionante construcción de su tiempo. Tampoco contábamos con la abigarrada profusión de huesos humanos y de animales que encontramos tras retirar los escombros de los primeros niveles. El suelo de la cámara principal aparecía cubierto por restos humanos desarticulados, desparramados con gran confusión, mezclados con fragmentos de cerámica y otros objetos. Se hizo necesaria la utilización de técnicas de decantación con todos los materiales procedentes de la tumba para extraer pequeños fragmentos de huesos y objetos varios, entre ellos huesos de aves y de peces.

El descombro de la cámara principal nos posibilitó el acceso a las cámaras laterales de la tumba. Cuatro de las seis se encontraban intactas. Después de limpiar los pasadizos descubrimos que las cámaras laterales no estaban bloqueadas; entrando a gatas, se accedía a una cámara de piedra perfectamente construida y hermosamente cubierta, que había existido durante más de 4000 años y donde podía permanecerse de pie. Más tarde, el carbono-14 reveló que la tumba se había utilizado desde el 3200 a. de C., aproximadamente; por tanto, la estructura tenía 5000 años: algunos siglos más antigua que las primeras pirámides de Egipto.

En la tumba había pocos objetos funerarios. Encontramos vestigios fragmentarios de 34 vasijas, tres cuchillos de piedra pulimentada y una pieza de sílex tallado. Los huesos de animales, claros indicios de restos de alimentos llevados al interior de la tumba, bien como ofrenda al difunto o como comida para consumir durante la celebración de los ritos funerarios, representaban, como mínimo, a siete ovejas, 18 corderos y cinco bueyes. También había huesos de ciervo común y otros mamíferos salvajes; asimismo se recogieron al menos 35 aves de varias especies y siete clases de peces.

Los análisis de los restos humanos, efectuados por Judson Chesterman, de la Universidad de Sheffield, ofrecieron resultados sorprendentes. De un total de 12.600 piezas que recuperamos, Judson Chesterman calculó que la cámara principal, parcialmente excavada, y una cámara lateral excavada del todo, aportaban datos de unos 157 individuos. Como en los túmulos alargados de Wessex, quedaba claro que el difunto había sido expuesto antes de reunir sus huesos para una posterior inhumación en la tumba. Si esta cifra se multiplica por un factor apropiado que considere el 20 por ciento de la cámara principal y las otras cinco cámaras laterales sin excavar, Quanterness quizá contenga los restos de no menos de 400 individuos.

La proporción existente entre los hombres y las mujeres (54:46) raya tanto la igualdad que rechaza una exclusión de la tumba por motivos del sexo. Excepto niños, aparecen todos los grupos de edad. Esta distribución por edades demuestra que la esperanza de vida se hallaba entre los 20 y 25 años. Aunque tal esperanza de vida es corta, no está en desacuerdo con los hallazgos de otros yacimientos neolíticos, y ello indica que los que nosotros descubrimos pudieron corresponder a los difuntos de una comunidad entera. De los datos aportados por el carbono-14 se deduce que la tumba estuvo en servicio unos 550 años, y por ello podemos establecer que el tamaño de la comunidad sería de unos 13 a 20 individuos.



6. TUMBA DE QUANTERNESS, que se muestra en este dibujo en sección tal y como debió aparecer su mitad septentrional poco después de la construcción. La figura humana del centro (*color*) nos da una idea de su tamaño; el perímetro exterior de piedra mide 45 metros; la altura en el centro de la tumba fue, originalmente, de 3,5 metros. Se aprecian tres de las seis cámaras laterales (*color claro*); el autor y sus colegas excavaron en su totalidad una de las cámaras laterales y parcialmente la principal. El cálculo global de enterramientos es de unos 400 individuos. (Recreación de Patricia J. Wynne.)

Mientras proseguían las excavaciones en Quanterness, examinamos otros yacimientos de las Orkneys. Los cálculos apuntaban en Quanterness hacia una inversión de trabajo de unas 7000 horas-hombre para la construcción de la tumba y su túmulo, si bien este cálculo no tiene en cuenta la destreza y habilidad de los constructores. La mayoría de las restantes tumbas cameradas de las Orkneys, ampliamente dispersas, hablaban de inversiones de trabajo del mismo orden.

Nuestra investigación en el otro *henge*, el anillo de Brogar, demostró que el foso que rodea al *henge* no se debía a ninguna remoción de tierra, sino al espléndido trabajo de corte de la piedra; supone una inversión de unas 80.000 horas de trabajo humano (y ello sin entrar en el cálculo la erección de las piedras clavadas alrededor). Desgraciadamente, no pudimos datar el monumento. La mejor aproximación a su cronología proviene del análisis por carbono-14 y lo sitúa en torno al 3000 a. C., fecha obtenida por Ritchie para su monumento más pequeño, el Stones de Stenness. Sí logramos obtener datos del foso que ciñe la gran tumba de Maes Howe; de acuerdo con los mismos, la estructura se construyó hacia el 2800 a. de C. Puesto que había que acarrear la piedra desde una considerable distancia, Maes Howe exigiría una inversión de trabajo mayor que la del anillo de Brogar y la de Quanterness.

Estos resultados nos permiten esbozar algunas conclusiones de carácter social sobre los primeros habitantes de las islas Orkney, que pueden cotejarse con las establecidas de modo más general para la arqueología social de Wessex. En primer lugar, en tiempos de la primera ocupación de la tumba de Quanterness, ésta no se encontraba estratificada espacialmente por arriba o por debajo; los monumentos se construyeron a diferente escala. No existía, pues, una jerarquización espacial. En segundo lugar, Quanterness fue una tumba de "acceso igualitario", teniendo en cuenta el equilibrio representativo en sexo y edad, exceptuados los niños. En tercer lugar, los objetos funerarios eran muy sencillos, lo que desvela la inexistencia de una jerarquización entre los componentes de la tumba. Este rasgo, junto con el acceso igualitario a la tumba, apoya la tesis de que los habitantes pertenecían a una sociedad bastante igualitaria. En cuarto lugar, las prácticas funerarias eran complejas: la inhumación de los huesos descarnados era el último estadio en el tratamiento del cadáver, habiéndole precedido la descomposición del mismo. Finalmente, la cámara estuvo en uso por lo menos cinco siglos.

Consideremos ahora otros hechos, deducidos de la naturaleza y de sucesivas vicisitudes del asentamiento. El grupo que utilizó el *cairn* no constaría probablemente de más de 20 individuos, incluyendo hombres, mujeres y niños; el trabajo necesario para la construcción del monumento, menos de 10.000 horas de trabajo humano, podría haber sido aportado sin dificultad por este grupo en un plazo de pocos años, o quizás en menos tiempo si los constructores contaban con la ayuda de sus vecinos; Quanterness es uno más de los numerosos *cairns* de tamaño similar que hay en las Orkney; la distribución de estos *cairns* es bastante dispersa, lo que indica que el grupo que los utilizaba ocupaba y explotaba el territorio circundante; finalmente, algunos siglos después de la construcción de Quanterness surgen, en el área central de la misma isla principal de las Orkney, unos cuantos monumentos construidos a mayor escala: el anillo de Brogar y Maes Howe. La nueva construcción implica, en cierto sentido, la aparición de una organización social más centralizada.

Los hallazgos procedentes de Quanterness y de yacimientos próximos nos autorizan a formular, por los detalles de la excavación, conclusiones que armonizan bien con las obtenidas en Wessex. Mientras que el tiempo transcurrido entre la constatación de sociedades paritarias y de sociedades centralizadas no es suficiente, en las Orkney, para demostrar de modo absoluto tal secuencia de desarrollo social, se ve claro que la aparición de los mayores monumentos, indicativa de un grado de centralización superior, fue relativamente tardía y, por ello, la mayoría de los *cairns* pertenecen a la fase más temprana, más igualitaria.

La abundancia de restos de esqueletos en Quanterness confirma la teoría de que los *cairns* no se construyeron para albergar sólo a uno o dos misioneros errantes. Sencillamente, los erigieron los habitantes del lugar, una población agrícola bastante dispersa. La habilidad de los constructores y cuidada elaboración de las estructuras apuntan al gran orgullo que el autor colectivo del monumento tendría de su obra.

No sostenemos aquí que los monumentos de las Orkney tuvieran un origen independiente de las otras áreas en las que se practicaban los enterramientos colectivos en tumbas monumentales. Ya sugerí más arriba que pudieron existir cinco o seis centros principales del megalitismo en Europa occidental, incluido uno en Irlanda y quizás otro más en el sur de Inglaterra. Es verosímil que la propagación de las técnicas agrícolas hacia Escocia se extendiera desde una u otra de estas zonas, y que los colonos trajeran consigo la técnica de la construcción de tumbas. Los primeros habitantes de las islas Orkney debieron pasar a las mismas desde el norte de Escocia; estaban familiarizados con las más simples tumbas megalíticas construidas por entonces en el norte. Lo que de especial tienen las



7. HUESOS de hombres y de animales se encontraron, mezclados, en el suelo de la cámara principal de la tumba de Quanterness. Escaseaban los objetos funerarios: unas cuantas piezas de sílex tallado, tres cuchillos de piedra pulimentada y fragmentos de unas 34 vasijas. Se calcula que la construcción de la tumba necesitó uns 7000 horas de trabajo, esfuerzo comparable a la construcción de un túmulo alargado.

Orkney es el impresionante desarrollo de la arquitectura local, basada en la excelente piedra arenisca del lugar, que produjo piezas maestras de la arquitectura prehistórica, como Quanterness y Maes Howe. Particular interés para el arqueólogo encierra la abundancia de restos prehistóricos de las Orkney, que hacen más fácil una interpretación coherente desde prespuestos sociales.

Las pruebas constatadas en las islas Orkney amplían significativamente el panorama obtenido en Wessex. Ello nos permite conjeturar sobre la función que pudieran desempeñar esos monumentos, al menos en alguna de las sociedades que los construyeron. Pero si se pretende explicar los orígenes del megalitismo, debe indicarse por qué se dio, cuándo se hizo y por qué precisamente en esos lugares.

Averiguar el "cuándo" no es lo más difícil. Los monumentos se remontan a los primeros agricultores de varias regiones de la Europa atlántica, al poco de su asentamiento y de haber adoptado un modo de vida estable a lo largo de las costas. Las poblaciones mesolíticas de cazadores, recolectores y pescadores, anteriores y más dispersas, que explotaban algunas de esas costas antes del desarrollo de la agricultura, no construyeron tumbas de tamañas proporciones.

"Dónde" resulta menos fácil. Si los pioneros de la agricultura europea hubieran desarrollado tales monumentos el problema sería menor. Pero, ¿por qué surgieron en estas particulares tierras costeras y no en el centro o este de Europa por ejemplo? Para contestar a esta pregunta debemos descubrir algún factor especial de la costa atlántica y demostrar su relevancia.

A tener en cuenta es la existencia de poblaciones mesolíticas dispersas, asentadas a lo largo de la costa atlántica, sobre todo en Portugal y la península bretona. Sabemos que estas gentes acostumbraban a enterrar los muertos en los concheros que iban abandonando. Sin embargo, nadie ha explicado todavía con exactitud por qué la existencia de estas primeras poblaciones facilitaría que los primeros agricultores construyesen tales monumentos.

Yo veo una vía de solución en un factor diferente, que puede expresarse en términos demográficos. Se ha dicho que el despliegue de la primera economía agrícola provocó un incremento sustancial de la población en todo el área que parte de Grecia y el Mediterráneo occidental y llega hasta un techo que culminó en lo que para ese tiempo fue una gran densidad de población. En cada caso, la aproximación a ese techo necesitaría de un declive proporcional al incremento. Tal declive no podría alcanzarse sin traba alguna. En Europa central siempre cabía aligerar el excedente de población con la emigración hacia el este y hacia el norte; de este modo podía evitarse el peligro de aproximarse al techo del desarrollo de la población. Por el contrario, en las tierras que bordean el Atlántico no cabía esa válvula de escape. Sencillamente, no había adónde emigrar y, por tanto, las consecuencias que se derivasen de este hecho habrían de afrontarse como mejor se pudiera.

Es en esas precisas circunstancias, diría yo, cuando la densidad de población se acercara a unas cotas peligrosas y urgiera la necesidad de contener el incremento de la población, resultarían apropiadas las decisiones que favoreciesen la buena cohesión social dentro de los pequeños grupos agrícolas y las vías de afirmación de sus derechos a cultivar la tierra. Las comunidades que habían erigido los monumentales hitos de demarcación del territorio disfrutarían de una ventaja de adaptación sobre grupos vecinos menos dogmáticos y con menor cohesión social. Las semillas de estas ideas quizá ya estuvieran presentes en las prácticas funerarias de las primeras poblaciones seminómadas. En cualquier caso, el incremento de la población, sin precedentes de tal magnitud, e imposible, por otra parte, de mitigar por medio de la emigración, haría válida cualquier decisión que potenciara la estabilidad y el orden social.

La explicación propuesta es sólo un esbozo. Necesita perfilarse con lo que nos ofrezcan ulteriores excavaciones. Sería conveniente, en este sentido, examinar los asentamientos donde los primeros agricultores pudieron entrar en contacto con los últimos supervivientes del Mesolítico. Se requieren más trabajos y más ideas para sostener que el problema se ha planteado en sus justos términos y que se ha encontrado el mecanismo para una solución satisfactoria.

Si el panorama aquí pergeñado puede aceptarse como plausible, ¿cabe sacar alguna conclusión más? ¿Puede probarse con el rigor propio de una teoría científica? En principio, los modelos espaciales pueden contrastarse a través de las modernas ténicas geográficas de análisis de localización. Pero en la práctica el nivel de comprobación no se aproxima al plano "isótropo" que asume frecuentemente el geógrafo; por otro lado, el paso del tiempo ha destruido muchos monumentos. Además, suele resultar difícil demostrar que los diferentes yacimientos estuvieron ocupados en un mismo momento, tal y como requiere un análisis estrictamente cuantitativo.

Nada prohíbe que hubiera más de un motivo para construir monumentos. La prehistoria nos ofrece notables concentraciones de tumbas encerradas en cementerios, no dispersas por tanto. Algunos de los argumentos utilizados aquí, que se fundan en la dispersión de las tumbas, no valdrían para los yacimientos funerarios cerrados. A pesar de estas limitaciones, el arqueólogo no debe desanimarse. Después de todo, la arqueología social trabaja bajo los mismos condicionantes que otras ciencias sociales. Se abordan actividades humanas que, en las circunstancias mejores, no son fáciles de cuantificar y resultan muy difíciles de prever. La mera posibilidad de identificar pautas sociales del pasado remoto, que sólo podemos situar en el tiempo valiéndonos del análisis por carbono-14, cuando se han perdido para siempre la mitad de las pruebas, es de suyo gratificante.

El historiador francés Fernand Braudel distinguía tres niveles en los que los procesos y los hechos, las "conjeturas" de la historia, se manifiestan por sí mismos. Los dramas personales, los "protagonistas" del pasado, son las crisis de corta duración. Subyacen a éstas los movimientos económicos y sociales que operan a plazo medio. Y, por debajo de éstos, se encuentra la secular tendencia, la larga duración, la *longue durée*, en la que intervienen factores de alcance mayor. Entre ellos, citemos las casi constantes limitaciones del entorno y las realidades básicas de la vida cotidiana. Curiosamente, el arqueólogo, que generalmente tiene poco que decir acerca del individuo o del acontecimiento de corta duración, puede a veces aproximarse más a los procesos de larga duración que el historiador de la modernidad, abrumado por el ruido y la profusión de testimonios del pasado reciente. Por esa misma razón, los arqueólogos que trabajan en temas que no se restringen a una cultura –piénsese en el origen de la agricultura, proceso de formación del estado o el tema aquí debatido de la construcción de los primeros monumentos del mundo, abordado desde una óptica social– saben que están aportando ideas muy fecundas que vale la pena examinar y contrastar, por duras que sean sus limitaciones prácticas.

Adaptación de un coleóptero a una semilla venenosa

Las larvas de un escarabajo brúquido se alimentan de semillas que contienen un aminoácido muy tóxico para otros insectos. Soslayan los efectos dañinos del aminoácido y aprovechan su nitrógeno

Gerald A. Rosenthal

La mayoría de los insectos se alimentan de vegetales. No es de extrañar, por tanto, que estos escarabajos se cuenten entre los adversarios más agresivos y destructores de las plantas. Aunque las plantas terrestres carecen de movilidad y no pueden huir para escapar de esas plagas destructoras, no son meros partícipes pasivos de la lucha que vienen librando insectos y vegetales desde hace más de 50 millones de años. Las armas principales de que se valen las plantas frente a los estragos alimentarios de los insectos son sus propios constituyentes químicos. Los miles de compuestos que sintetizan las plantas superiores se encuadran en dos categorías: metabolitos primarios y metabolitos secundarios. Los primarios son sustancias del tipo de los ácidos nucleicos, el adenosín trifosfato (ATP) y la glucosa, comunes a todos los seres vivos. Los secundarios, que difieren de una planta a otra, son esenciales en la guerra química que vienen sosteniendo.

Muchos biólogos consideran que la actividad alimentaria de los insectos y otros herbívoros, sumada a los estragos de una hueste de organismos patógenos, han proporcionado la presión de selección necesaria para la elaboración y conservación de refinadas barreras químicas contra la depredación y la enfermedad. En tal situación, la selección natural habría explotado la variabilidad intrínseca de la química secundaria de los vegetales para favorecer la elaboración de metabolitos que aumentaran la eficacia biológica del organismo (en la acepción darwinista de la expresión), proporcionándole un nivel seguro de protección. Ante la relación que se aprecia entre determinados insectos y vegetales, Paul P. Feeny, de la Universidad de Cornell, ha comentado: "Así, estamos presenciando una carrera de armamentos evolutiva en la que las plantas, para su supervivencia, han de asignar una fracción de su presupuesto metabólico a la defensa (tanto física como química) y los insectos han de dedicar una cuota de su energía y de sus nutrientes asimilados a varios dispositivos para la detección y el ataque de la planta explotada."

Este aspecto se ilustra claramente en la planta *Dioclea megacarpa*, una leguminosa parecida a una enredadera. Tiene un único insecto depredador: *Caryedes brasiliensis*, un pequeño escarabajo de la familia Brúquidos, coleópteros comedores de semillas de distribución mundial. El éxito de la planta en su repulsión de otros insectos se debe en gran parte a que almacena L-canavanina, un aminoácido insecticida que, entre otras propiedades, desbarata la producción de las proteínas normales de los insectos. Más notable aún resulta la adaptación de *C. brasiliensis* a la defensa química de la planta.

Nuestro trabajo en la Universidad de Kentucky ha vertido alguna luz sobre las interacciones que se dan entre los insectos y los metabolitos secundarios de las plantas superiores. La canavanina es uno de los aproximadamente 250 aminoácidos que sintetizan las plantas superiores y que no se utilizan como sillares de las proteínas. Este aminoácido no proteínico lo sintetizan los miembros de las Lotoideas (Fabáceas), un grupo de plantas leguminosas; se encuentra en cultivos de interés agronómico, como la alfalfa y numerosos tréboles, entre ellos los del género *Trifolium*. También lo sintetizan especies anuales y perennes, ornamentales y arborescentes, como *Wisteria* y *Robinia*. Existen indicios suficientes de que la canavanina constituye un importante metabolito de reserva de nitrógeno, en particular en la semilla, donde sostiene el crecimiento de la planta que empieza a desarrollarse.

La estructura de la molécula de L-canavanina es análoga a la de la L-arginina, uno de los 20 aminoácidos formadores de proteínas. En la canavanina, el grupo terminal $-CH_2$ de la arginina aparece sustituido por oxígeno. Tal diferencia es de escasa importancia en el metabolismo de la canavanina, que actúa prácticamente en todas las reacciones controladas por enzimas donde la arginina es el sustrato preferido. Por ejemplo, la canavanina se activa por acción de la arginil-ARNt-sintetasa, un enzima que activa la arginina y luego la enlaza a su ARN de transferencia. Si se ha unido canavanina al ARN de transferencia que normalmente transporta arginina a los puntos de ensamblaje de proteínas ubicados en los ribosomas, inevitablemente la canavanina se insertará en la cadena proteica en crecimiento en el lugar que correspondía a la arginina.

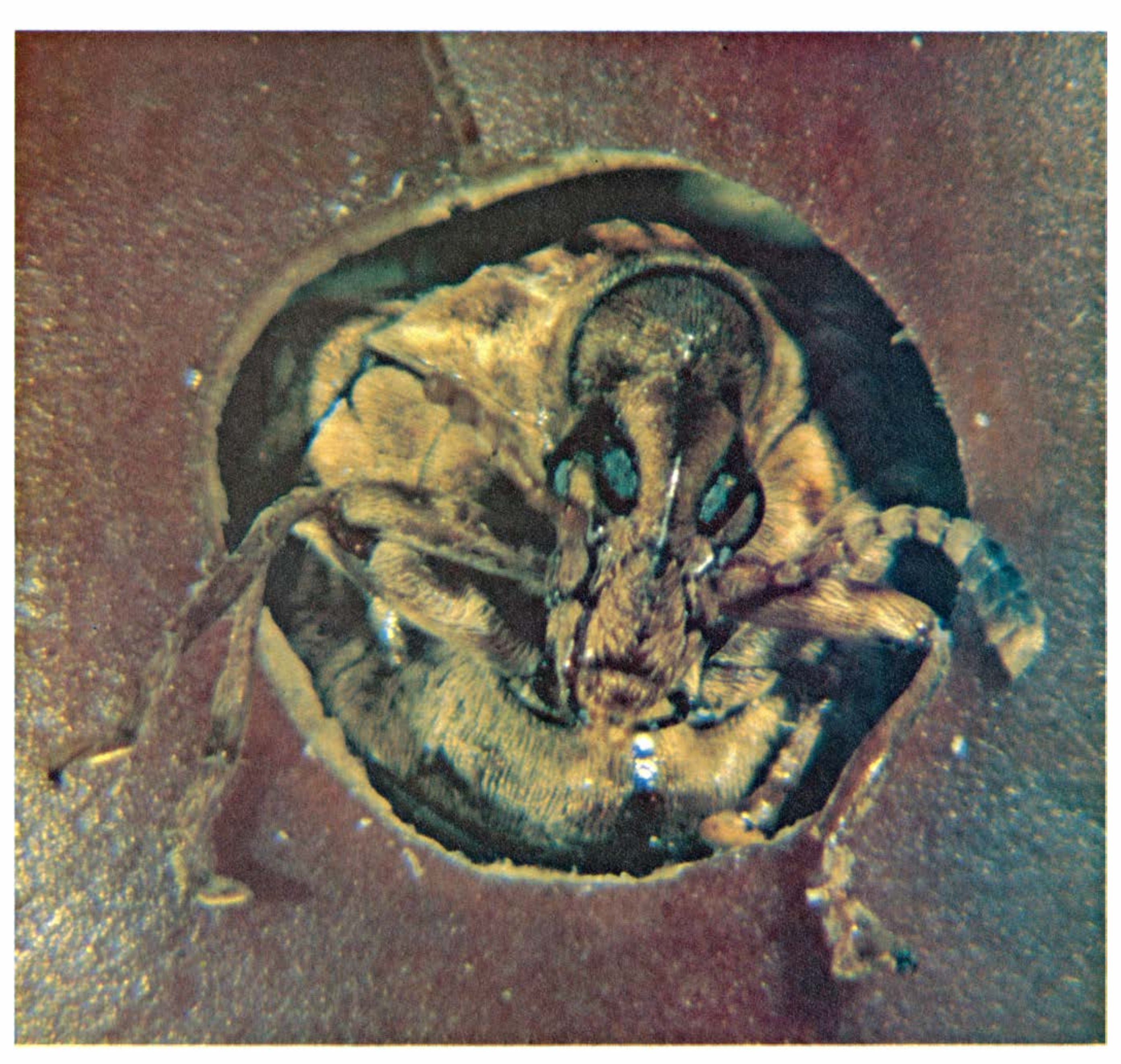
Desde el punto de vista químico, la canavanina no es tan básica como la arginina y, en condiciones fisiológicas, su carga positiva será inferior que la de la arginina. Esta diferencia de carga afecta a las interacciones que permiten el plegamiento de la cadena proteica en una conformación correcta única. Parece confirmarse que las proteínas con alteraciones estructurales causadas por la incorporación de canavanina, en sustitución de arginina, no funcionan adecuadamente. Ahí reside un importante efecto tóxico de la canavanina sobre la mayoría de insectos.

Con mi colega Douglas L. Dahlman hemos establecido experimentalmente las potentes propiedades insecti-

cidas de la canavanina. En nuestro trabajo utilizamos la esfinge del tabaco, *Manduca sexta*, un lepidóptero que consume plantas del tipo del tabaco y el tomate, ninguna de las cuales almacena canavanina. Se tenía así la seguridad de que *M. sexta* no estaba adaptado a ese aminoácido tóxico ni podía anularlo. Cuando inyectamos canavanina marcada con carbón radiactivo (C-14) en la hemolinfa, o líquido circulatorio, de la esfinge del tabaco, descubrimos que al menos el 3,5 por ciento de la canavanina marcada se incorporaba en proteínas recién formadas en el insecto.

Si se proporciona canavanina al cuerpo graso, un órgano interno, de la langosta migradora *Locusta migratoria migratorioides*, se detecta la incorporación del aminoácido en la vitelogenina, la principal proteína del cuerpo graso, muy importante en el desarrollo del huevo. La vitelogenina de la langosta que contiene canavanina presenta, en un campo eléctrico, más movilidad que la proteína natural. Este hallazgo revela que la incorporación de canavanina provoca un cambio significativo en las propiedades físico-químicas de la proteína. La canavanina se asimila exclusivamente a expensas del contenido en arginina de la proteína.

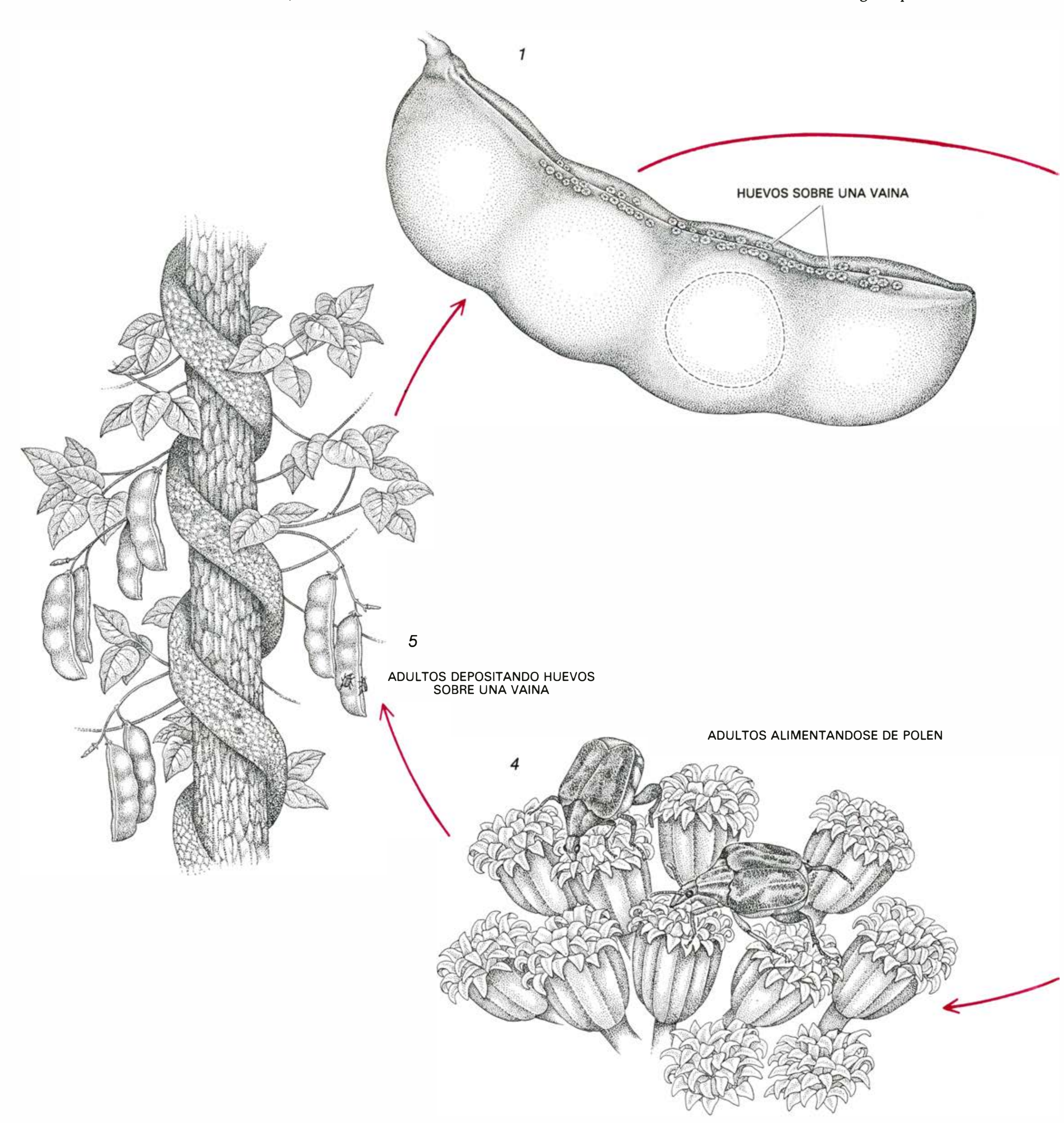
En 1973 conocí el trabajo de Daniel H. Janzen, un ecólogo tropical que en la actualidad se encuentra en la Universidad de Pennsylvania. Parte de la labor de Janzen en Costa Rica tenía que ver con estudios de la interacción del escarabajo brúquido *Caryedes brasiliensis* con la leguminosa *Dioclea megacarpa*, parecida a una enredadera. Mi interés por esta interacción concreta entre insecto y planta nació del descubrimiento de Janzen de que las semillas de la planta contenían considerable cantidad de canavanina. (Posteriormente determiné que la canavanina suponía hasta el 13 por ciento del peso



1. ESCARABAJO ADULTO de la familia Brúquidos, saliendo de una semilla de la leguminosa *Dioclea megacarpa* tras haber permanecido varios meses en su interior durante los estadios larvario y de ninfa. El escarabajo (*Caryedes brasiliensis*) ha recortado un agujero en la cubierta de la semilla para salir. La planta sintetiza el aminoácido canavanina, que es tóxico para otros insectos. *C. brasiliensis* es su único insecto depredador. A través de varias adaptaciones bioquímicas, las larvas (de las que puede haber hasta 50 en una misma semilla) evitan el efecto tóxico y aprovechan el nitrógeno de la canavanina.

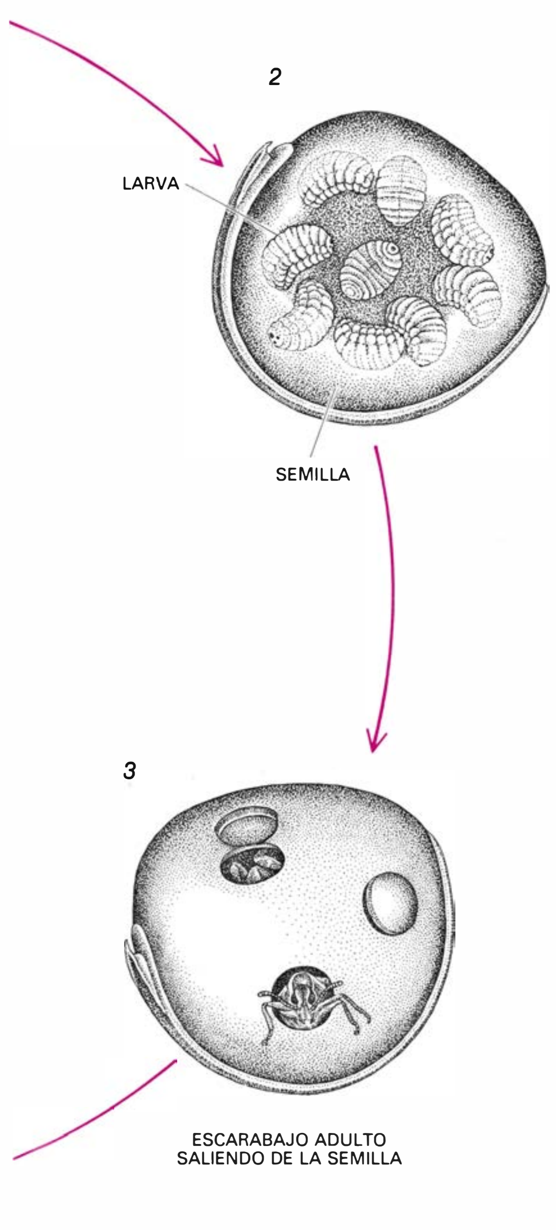
seco de la semilla, excluyendo la cubierta de la misma, y hasta el 55 por ciento de todo el nitrógeno de la semilla.) Tal acumulación de canavanina proporciona una barrera química muy eficaz frente a la depredación: los únicos consumidores de las semillas, entre los insectos, son las larvas de *C. brasiliensis*.

La hembra del escarabajo deposita sus huevos sobre la cubierta del fruto que acaba de madurar, a menudo cerca de la fisura característica de las leguminosas. Los huevos se hallan protegidos en estructuras denominadas ootecas. La puesta suele efectuarse a finales de otoño, cuando termina la estación de las lluvias y alrededor de un mes antes de que las cápsulas maduren. Las vainas de *D. megacarpa* sólo son



2. CICLO BIOLOGICO de una nueva generación del escarabajo brúquido *C. brasiliensis*; se inicia cuando la hembra adulta deposita sus huevos sobre una vaina de la leguminosa *D. megacarpa*, parecida a una enredadera. Los huevos se ponen a finales de otoño. Cuando las larvas emergen, perforan la pared de la vaina y la envuelta de una semilla, instalando su residencia en los tejidos de ésta. Meses después, las larvas se han transformado en ninfas y luego en adultos; los jóvenes escarabajos recortan puertas en el recubrimiento de la semilla, salen a la vaina y de allí al exterior, a través de la sutura de la

adecuadas para la deposición de los huevos durante unos pocos meses. Los coleópteros brúquidos no ponen sus huevos sobre vainas maduras o en descomposición; si las hembras no aciertan el período adecuado, la población de escarabajos debe esperar a que termine



misma o de una abertura practicada por otro animal (que no puede ser insecto). Los escarabajos de la especie se alimentan del polen de varias plantas; completan el ciclo poniendo huevos en otoño.

la siguiente estación de lluvias para completar el ciclo biológico con la oviposición en la cosecha de vainas siguiente.

Tras su eclosión, la larva perfora la pared del fruto y de la envoltura de la semilla y construye una pequeña cámara en los tejidos de reserva de la semilla. La cámara será el hogar larvario durante los meses que precisa la larva para crecer, pupar y emerger transformada en adulto. Hasta 50 escarabajos pueden desarrollarse en una misma semilla. El joven escarabajo perfora un boquete de salida en la envoltura de la semilla y escapa por la sutura que corre a lo largo de la legumbre, o bien a través de una abertura producida por un vertebrado que abra la legumbre para acceder a sus semillas. El escarabajo adulto, que emerge de abril a junio, se alimenta de polen; no ingiere canavanina.

Este depredador de semillas constituye un sistema ideal para investigar de qué modo los insectos se adaptan a los tóxicos que producen las plantas. La capacidad del escarabajo para llevar a término con tanto éxito su hábito alimentario especializado se apoya en determinadas adaptaciones bioquímicas que le permiten neutralizar la canavanina, que de otro modo le sería altamente venenosa. Formando parte de un programa de investigación que la National Science Foundation patrocinó y yo dirigí, emprendí una investigación sistemática y a largo plazo de estas adaptaciones bioquímicas.

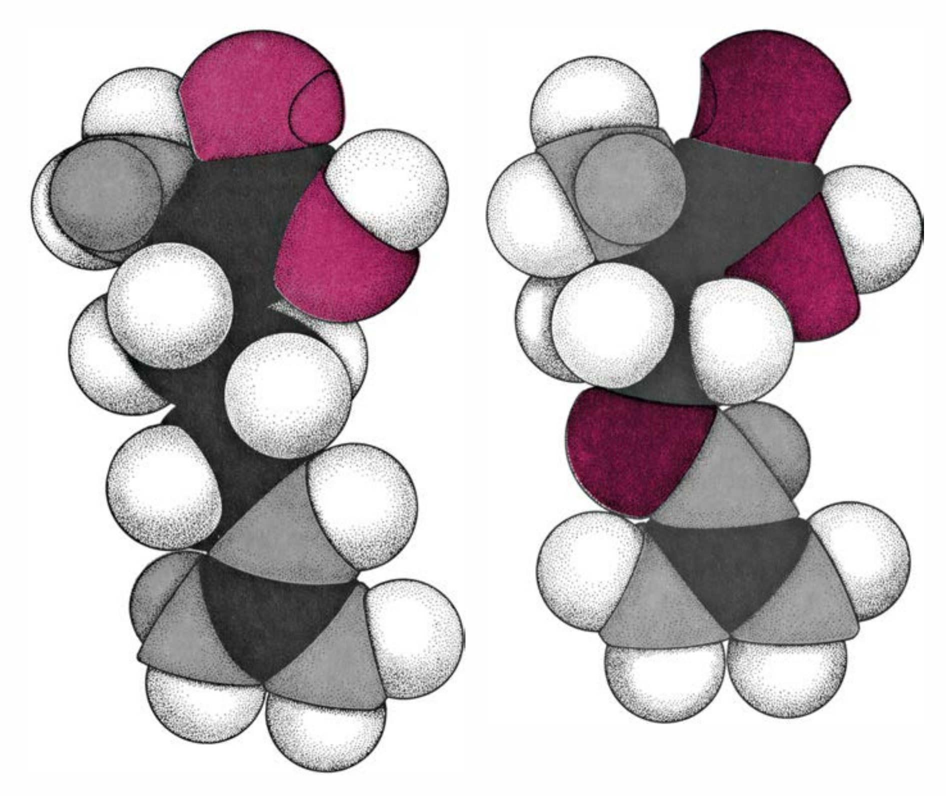
Había que empezar por averiguar si este escarabajo brúquido incorporaba o no canavanina a sus proteínas. En caso afirmativo, ¿cómo evitaba la síntesis de proteínas aberrantes? Iniciamos los trabajos inyectando canavanina marcada radiactivamente a larvas obtenidas de semillas infestadas de *D. megacarpa*. Sólo se marcaba el átomo de carbono terminal de la canavanina inyectada. Esta pauta única de marcado era importante, porque el tratamiento de la canavanina radiactiva con los enzimas arginasa y ureasa libera el átomo de carbono radiactivo en forma de dióxido de carbono, gas que puede atraparse químicamente y luego cuantificarse mediante espectroscopía de centelleo en fase líquida.

Se aislaron las proteínas recién sintetizadas y ligeramente radiactivas extraídas de las larvas y se digirieron con un ácido fuerte para soltar sus aminoácidos constituyentes. El digerido proteico se purificó luego mediante cromatografía de intercambio iónico, que aislaba aminoácidos básicos, como la canavanina. El tratamiento de los aminoácidos aislados con ureasa y arginasa no produjo la liberación de cantidades apreciables de dióxido de carbono radiactivo. Este hallazgo demostraba que las larvas no incorporaban canavanina en sus proteínas en cantidad significativa, aunque se había dado una cierta inserción del átomo de carbono radiactivo de la canavanina en los aminoácidos proteicos.

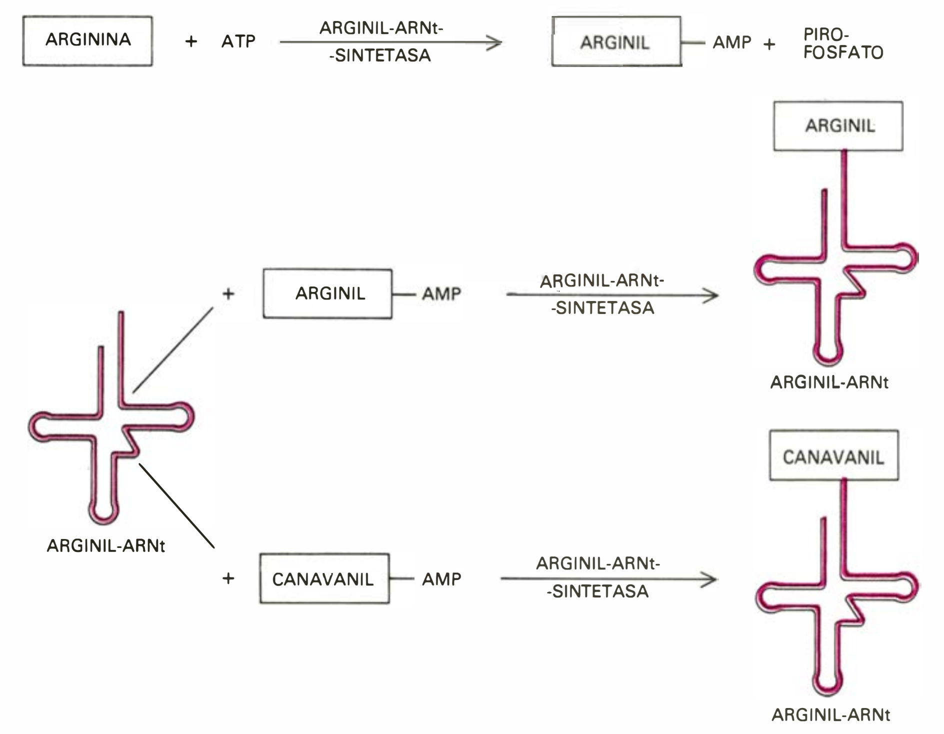
Se puso de manifiesto la base bioquímica del hallazgo al comparar la arginil-ARNt-sintetasa de este insecto con la de *Manduca sexta*, la esfinge del tabaco. El enzima de *M. sexta* unía la arginina y la canavanina al arginil-ARNt, pero el enzima activador de la arginina de *C. brasiliensis* sólo unía la arginina. Así, parecía probable que entre las adaptaciones bioquímicas conseguidas por el escarabajo se hallara el desarrollo de una sintetasa arginil-ARNt, capaz de distinguir la canavanina de la arginina. La canavanina no se unía al ARN de transferencia para la arginina y el coleóptero evitaba la producción de proteínas aberrantes.

En un reciente estudio paralelo precisamos algo más la capacidad discriminadora del sistema de síntesis de proteínas del coleóptero brúquido. El estudio requirió la síntesis de un grupo de aminoácidos marcados radiactivamente, estructuralmente relacionados con la arginina, así como su examen, junto con la canavanina y la arginina, para determinar si alguno de ellos se incorporaba en las proteínas de *M. sexta* y de *C. brasiliensis*. Las larvas de la esfinge del tabaco fijaban todos los compuestos a proteína recién sintetizada; no así las larvas del brúquido. La notoria capacidad de este animal depredador de semillas para distinguir entre la arginina y otras moléculas estructuralmente parecidas a ella le salva de incorporar aminoácidos incorrectos a las proteínas.

Quedaban muchas cuestiones importantes sin resolver. Centramos nuestra atención en el procedimiento del que se valía la larva del coleóptero para neutralizar la canavanina. ¿Excreta la sustancia o evita sus efectos bioquímicos adversos por otra vía? Cabía otra posibilidad muy sugestiva: que el escarabajo hubiera desarrollado algún medio de utilizar el producto químico tóxico natural como recurso alimentario. Era razonable pensar que el insecto no iba a desperdiciar una fuente tan abundante de nitrógeno.



3. SEMEJANZA entre las estructuras moleculares de la L-arginina (*izquierda*) y L-canavanina (*derecha*). Los átomos constituyentes son carbono (*negro*), hidrógeno (*blanco*), nitrógeno (*gris*) y oxígeno (*color*). La arginina es un aminoácido de las proteínas; la canavanina es uno de los 250 aminoácidos que sintetizan las plantas superiores y no se utilizan para ensamblar proteínas. La diferencia estructural entre los dos aminoácidos es pequeña: el grupo metileno (-CH_2) terminal de la arginina aparece sustituido por oxígeno en la canavanina. Además, el grupo guanidino de la arginina se encuentra en la forma imina (=NH), mientras que el grupo correspondiente en la canavanina se encuentra en forma amina (−NH_2).



4. PAPEL DEL ARN DE TRANSFERENCIA al asociarse con la arginina o, equivocadamente, con la canavanina en la elaboración de proteínas. Normalmente (*arriba*) el enzima arginil-ARNt-sintetasa reacciona con la arginina en presencia de adenosín trifosfato (ATP) y produce una forma activada del aminoácido. El enzima interviene también en el enlace de la arginina activada con el ARN de transferencia correcto. Por lo general (*centro*) la arginil-ARNt-sintetasa une la arginina a su ARN de transferencia antes de entregar la arginina a un complejo ribosómico, donde se fabrica la proteína. Si en lugar de arginina se une canavanina (*abajo*), el ARN de transferencia que normalmente transporta arginina la introduce en el sitio de emsamblaje. La síntesis de proteína no dispone de ningún mecanismo que efectúe la verificación del aminoácido fijado al ARNt, con lo que se incorpora canavanina en la proteína.

Nuestro primer esfuerzo se dirigió a analizar el contenido en canavanina de las semillas de *D. megacarpa* infestadas con larvas del escarabajo brúquido. Aplicamos una técnica estadística para estimar el peso de la semilla antes del ataque del insecto a partir del peso de la envuelta intacta de la semilla. La cantidad de canavanina contenida en la parte no comida de la semilla y en el material fecal de las larvas nos permitía calcular cuánta canavanina debía haber originariamente en la semilla. Resultaba que más de la mitad de la canavanina original de la semilla se consumía en la alimentación de las larvas.

Quedaba por averiguar cómo se metabolizaba la canavanina ingerida. La arginasa, un enzima que separa la L-arginina en L-ornitina y urea, actúa también sobre la L-canavanina, escindiéndola en L-canalina y urea. Este enzima se halla ampliamente distribuido en los insectos y desempeña varias funciones importantes, siendo una de ellas la de proporcionar ornitina a partir de la arginina para la formación de ácido glutámico. El escarabajo brúquido produce arginasa, por lo que podría convertir la canavanina en canalina y urea: $H_2N(C\text{-}NH_2) = N\text{-}O\text{-}CH_2\text{-}CH_2\text{-}CH(NH_2)\text{-}COOH$ [L-canavanina] → $H_2N\text{-}O\text{-}CH_2\text{-}CH_2\text{-}CH(NH_2)COOH$ [L-canalina] + $H_2N\text{-}C(=O)NH_2$ [urea].

La ureasa, el enzima que actúa sobre la urea, se encuentra raramente en los insectos. Nuestro análisis de las larvas del coleóptero brúquido reveló una extraordinaria actividad ureásica. Las larvas escinden la canavanina en canalina y urea; luego, recurren a la ureasa para generar amoníaco y urea. De este modo, la mitad del nitrógeno almacenado en la canavanina puede incorporarse al metabolismo en forma de amoníaco.

¿Utiliza la larva el amoníaco como fuente de nitrógeno para sus reacciones metabólicas? Además de detoxificar la canavanina, ¿incorpora la larva parte del nitrógeno de este compuesto en los aminoácidos que sintetiza? Se trata de preguntas importantes porque, con la posible excepción de la incorporación de cianuro en la L-asparagina, un aminoácido de las proteínas, no se conoce otro ejemplo de insecto que aproveche un compuesto vegetal tóxico para fabricar aminoácidos.

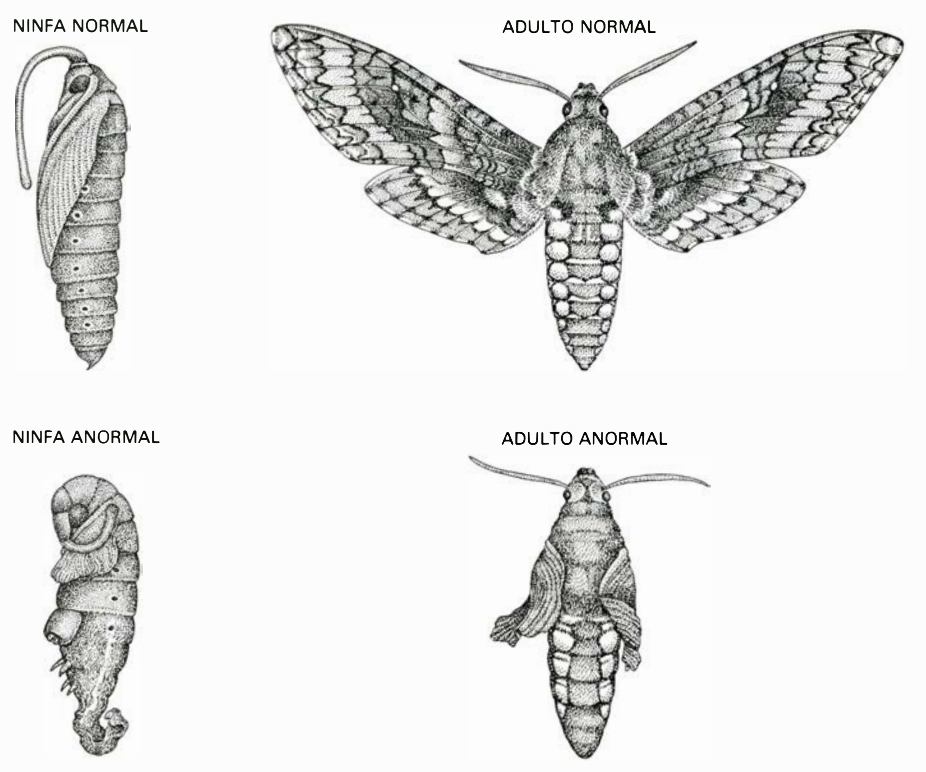
La disponibilidad comercial de urea marcada con el isótopo pesado del nitrógeno (N-15) y la facilidad con la que la larva del brúquido convierte la canavanina en urea permitió comprobar experimentalmente esta hipótesis. Se inyectó a las larvas urea con el isóto-

po pesado de nitrógeno; se aislaron los aminoácidos libres acabados de sintetizar y se convirtieron químicamente en una forma más volátil. Esta conversión facilitó la separación de los aminoácidos mediante cromatografía de gases. A medida que los compuestos salían del aparato cromatográfico se les introducía en un espectrómetro de masas. (Estos experimentos se realizaron en las excelentes instalaciones de espectroscopía de masas del Instituto de Investigación del Tabaco y la Salud, de la Universidad de Kentucky.)

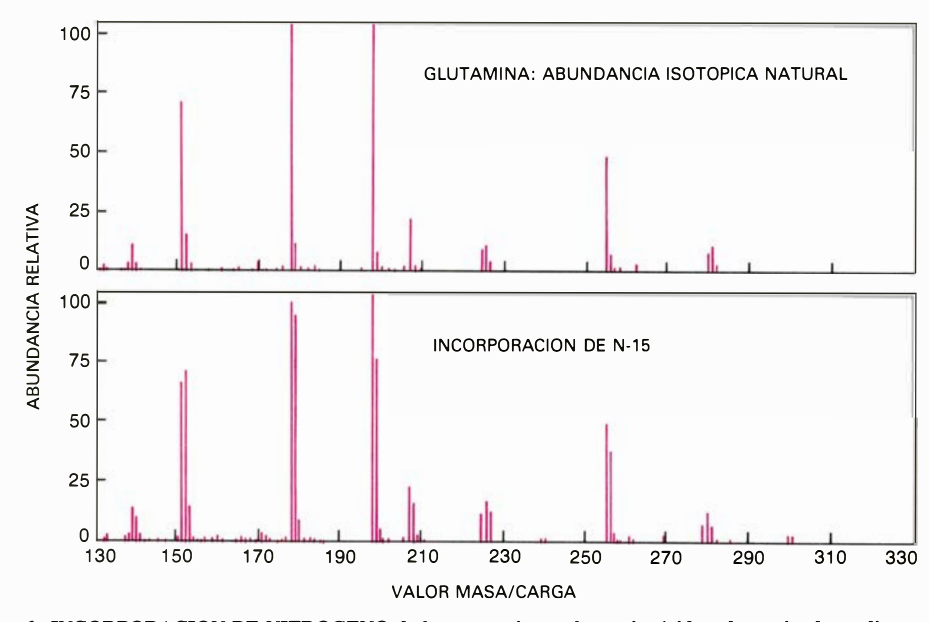
En el espectrómetro de masas el compuesto se bombardea con electrones que rompen el aminoácido en un modelo característico de fragmentos. El análisis mediante ordenador proporcionó una caracterización y compilación exactas de los fragmentos obtenidos para cada compuesto. Conociendo la abundancia relativa de los fragmentos que contienen los dos isótopos de nitrógeno puede reconstruirse con precisión las proporciones de N-15 y de N-14 del aminoácido original.

Estas determinaciones revelaron una incorporación significativa de nitrógeno pesado a los aminoácidos alanina, glicina, serina, prolina, metionina, ácido aspártico (o asparagina) y ácido glutámico (o glutamina). Se trata de los aminoácidos formadores de proteínas que los insectos pueden sintetizar a partir de precursores adecuados. En cambio, no se encontró nitrógeno pesado apreciable en la treonina, leucina, isoleucina, histidina, lisina o hidroxiprolina. Son éstos los aminoácidos formadores de proteínas que los insectos no pueden sintetizar y deben obtener de su alimento. Se cree asimismo que la valina es esencial en la dieta de los insectos, pero *C. brasiliensis* podía sintetizar cantidades muy pequeñas de la misma. Por limitaciones técnicas no se evaluaron la tirosina, fenilalanina, arginina y triptófano. Resultó interesante constatar que la larva del brúquido incorporaba también cantidades apreciables de N-15 en varios aminoácidos no proteínicos: ácido 2-aminobutírico, 3-alanina y homoserina.

¿Cómo sostenía la fabricación de estos áminoácidos el nitrógeno de la canavanina, que, en última instancia, se libera en forma de amoníaco? Pudiera ser que el ácido glutámico y la glutamina desempeñasen un papel importante en el metabolismo de aminoácidos del insecto. Entre los insectos se distribuyen ampliamente los enzimas que pueden catalizar la síntesis de alanina y ácido aspártico a partir del ácido glutámico. La glutamina y la asparagina se obtienen fácilmente a partir, respectivamente, del ácido glutámico y del ácido aspártico, mediante la simple adición de amoníaco a estos aminoácidos. El ácido glutámico se halla asimismo directamente ligado a la síntesis de prolina y serina; la glicina se forma a partir de serina. La considerable cantidad de nitrógeno pesado que se encuentra en el ácido 2-aminobutírico puede parecer inesperada, pero el compuesto se forma a partir de treonina y metionina. Así, prácticamente todos los aminoácidos de los que se tiene noticia que sintetizan los insectos, incluidos los que



5. EFECTOS DE LA CANAVANINA en *Manduca sexta*, la esfinge del tabaco, que no suele alimentarse de plantas con canavanina. Las formas anormales de ninfa y adulto se deben a la inclusión de canavanina en la dieta que, muchos días antes, se suministra en el laboratorio a las larvas. Se provocan anomalías de cierta gravedad inyectando directamente canavanina en la hemolinfa, el fluido circulatorio del insecto.



6. INCORPORACION DE NITROGENO de la canavanina en los aminoácidos, determinada mediante espectroscopía de masas. A larvas de *Caryedes brasiliensis* se les inyectó urea marcada con el isótopo pesado del nitrógeno (N-15). Los aminoácidos que sintetizaron se separaron mediante cromatografía de gases y se analizaron con un espectrómetro de masas los diversos compuestos. Los resultados demuestran que las larvas incorporaron N-15 en los aminoácidos proteínicos que pueden sintetizar, pero no en los que deben procurarse de la dieta. Uno de los aminoácidos formadores de proteínas que incorporó cantidades significativas de N-15 fue la L-glutamina; se representan aquí sus datos de espectrometría de masas.

sintetiza *C. brasiliensis*, encierran cantidades apreciables de nitrógeno equivalentes a la cantidad suministrada por la canavanina.

Surge un nuevo problema: la L-canalina, un aminoácido no proteico producido por la escisión de la L-canavanina en urea, es también muy tóxica. Se distingue por ser el único aminoácido no proteínico con un grupo aminooxi terminal libre: -ONH_2. Experimentos con larvas de la esfinge del tabaco revelaron que la L-canalina retardaba el crecimiento, producía graves aberraciones del desarrollo, aumentaba la mortalidad y obstruía la función nerviosa. La canalina reacciona con el grupo aldehído del fosfato de piridoxal, un cofactor esencial para la función de determinados enzimas. El complejo canalina-fosfato de piridoxal es estable, y su formación suprime esencialmente la acción catalítica de determinados enzimas que contienen fosfato de piridoxal. A primera vista, parece que el coleóptero brúquido se limita a sustituir un veneno por otro.

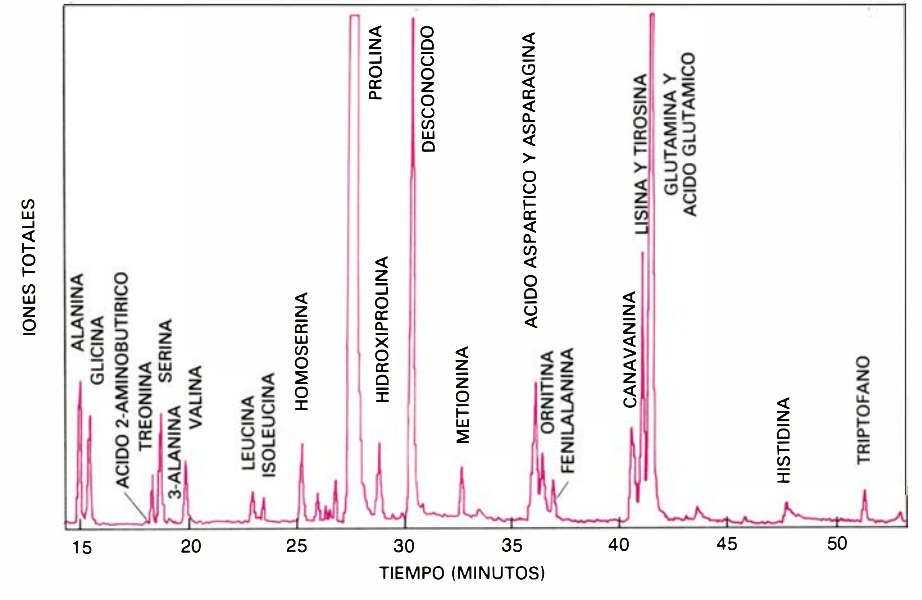
Al examinar el destino metabólico de la canalina en las larvas del brúquido encontramos que el insecto dispone de un enzima capaz de escindir la canalina para producir homoserina: H_2N-O-CH_2-CH_2-$CH(NH_2)COOH$ [L-canalina] → HO-CH_2-CH_2-$CH(NH_2)$-COOH [L-homoserina] + NH_3. Se plantea entonces la cuestión de cómo se las apaña el insecto con todo el amoníaco, que es una forma tóxica de nitrógeno reducido. El principal producto nitrogenado de excreción en los insectos es el ácido úrico; el amoníaco suele ser un producto de excreción menos importante. Analizando las heces de las larvas, encontramos que, en ellas, cerca del 90 por ciento del nitrógeno se hallaba en forma de amoníaco y urea. Sólo el 11 por ciento del nitrógeno total excretado se hallaba en forma de ácido úrico. Es evidente que este brúquido que se alimenta de semillas excreta parte del amoníaco ajeno directamente en las heces.

La eliminación de amoníaco a través de las heces no es el único mecanismo de que dispone el insecto para evitar los efectos tóxicos de su exceso. La larva del escarabajo posee una glutamina sintetasa muy activa; se trata de un enzima que añade amoníaco al ácido glutámico para formar glutamina. La glutamina es un importante aminoácido libre de la hemolinfa del insecto. El análisis de los aminoácidos de la hemolinfa revela asimismo un gran pico para la prolina, aminoácido que quizá se emplee como fuente de carbono esquelético para la formación de ácido glutámico antes de la constitución de glutamina. La secuencia de la reacción desde prolina hasta ácido glutámico y de éste a glutamina puede revelar un mecanismo interesante para canalizar amoníaco tóxico hacia una forma inocua.

En lo concerniente a la homoserina, las investigaciones de John Giovanelli y sus colegas, del Instituto Nacional de Salud Mental, han establecido que las plantas emplean derivados homoserínicos como precursores en la fabricación de determinados aminoácidos indispensables que contienen azufre. Cabe esperar que muchos tipos de insectos (por ejemplo, los pulgones que se alimentan de las semillas del guisante, *Pisum sativum*) utilicen homoserina, porque este aminoácido no proteínico puede alcanzar más del 10 por ciento del peso seco de ciertas partes de la planta.

Después de administrar homoserina marcada con carbono 14 a las larvas del escarabajo brúquido, observamos una transferencia detectable, aunque insignificante desde el punto de vista cuantitativo, de los átomos radiactivos a los aminoácidos metionina y cisteína. Análisis automatizados de los aminoácidos del extracto de estos insectos tratados con homoserina revelaron un gran pico de radiactividad que no correspondía a ningún aminoácido. Lo confirmó el hecho de que la sustancia desconocida no formaba complejo coloreado con la ninhidrina. Este reactivo reacciona con los aminoácidos y confiere un característico color azul púrpura; es el clásico ensayo colorimétrico de los aminoácidos.

Quizás el escarabajo brúquido utilice la homoserina para obtener nitrógeno destinado a la síntesis de otros aminoácidos, a través de la acción de una transaminasa. La reacción de este enzima con la homoserina haría que ésta perdiera su respuesta positiva normal a la ninhidrina. En cualquier caso, no cabe duda de que el insecto sintetiza cantidades considerables de homoserina que no emplea en la fabricación de aminoácidos azufrados.



7. RESULTADOS DEL ANALISIS por ordenador de la incorporación de nitrógeno pesado a los aminoácidos de *C. brasiliensis*. Demuestran que el escarabajo brúquido utiliza el nitrógeno de la canavanina, liberado por la formación y degradación de urea (donde intervienen la arginasa y la ureasa), en la síntesis de los aminoácidos que puede fabricar, pero no en la de los que el insecto obtiene en su dieta normal.

Un aspecto fascinante de estos procesos bioquímicos es que el brúquido se ha adaptado a una sustancia perjudicial de su dieta (la canavanina) comportándose cual si fuera una planta sintetizadora de esa sustancia. Esta evolución convergente la apoyan cuatro tipos de pruebas. La primera es la formación de un enzima: arginil-ARNt-sintetasa, capaz de discriminar entre arginina y canavanina. La capacidad de distinguir entre un aminoácido que normalmente forma parte de una proteína y otro que no es un mecanismo conocido, por cuyo medio las plantas evitan los efectos adversos de sus propios aminoácidos no proteínicos.

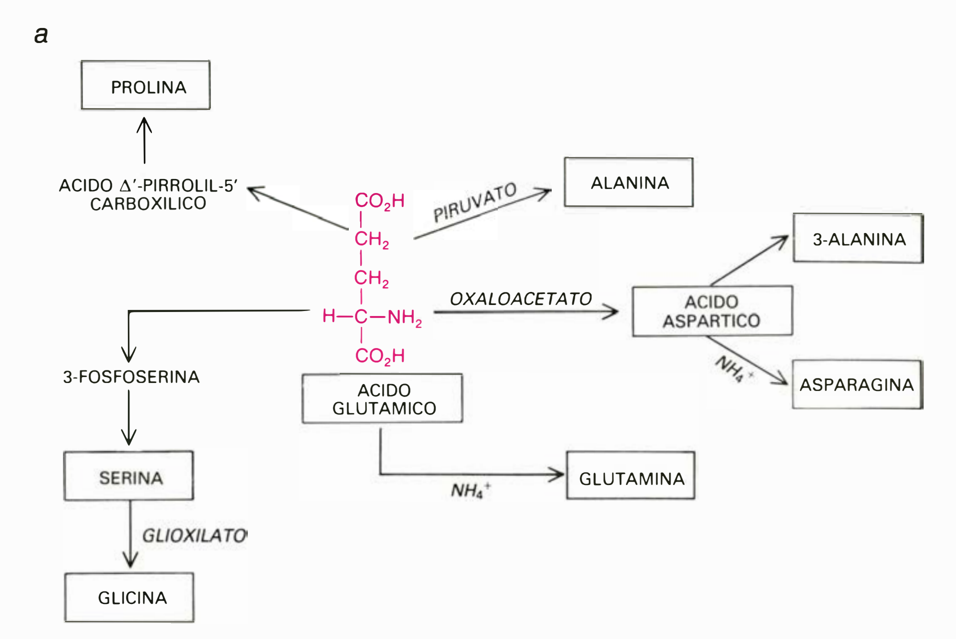
La segunda serie de indicios es que la formación de urea a partir de canavanina, unida a la formación de amoníaco desde la urea, es el mecanismo principal por el que las plantas que producen

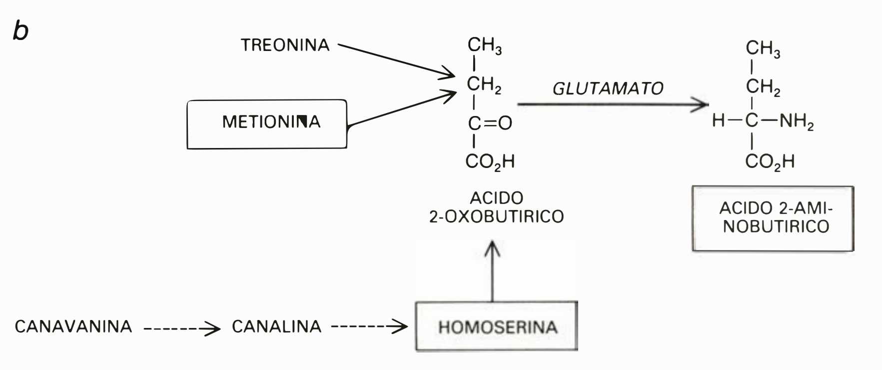
canavanina movilizan el nitrógeno almacenado en ella. La tercera es que el insecto detoxifica la canalina al convertirla en homoserina; en las plantas que sintetizan canavanina la conversión de homoserina en canalina puede formar parte de la síntesis de canavanina.

Hallamos, por último, que la arginasa de *C. brasiliensis* tiene una afinidad aparente mayor por la canavanina, y reacciona más rápidamente con la canavanina como sustrato (en comparación con la arginina), que las arginasas de otros insectos que probamos. Una distinción similar se encuentra al comparar la arginasa de *Canavalia ensiformis*, la judía sable, una leguminosa que utiliza canavanina, con la arginasa de la soja (*Glycine max*), que no sintetiza canavanina.

Otra conclusión que se puede extraer de nuestras investigaciones es que este escarabajo brúquido comedor de semillas ha desarrollado muchas adaptaciones bioquímicas distintas, aunque interrelacionadas, en su explotación de *D. megacarpa* con fines alimentarios. Es razonable proponer que entre las primeras adaptaciones se cuenta la capacidad discriminadora que le permite evitar la síntesis de proteínas no funcionales. Quizá los brúquidos que invadieron originariamente la leguminosa se limitaran a excretar la canavanina y la canalina, si no a neutralizarla. Otros metabolitos ricos en nitrógeno almacenados en la semilla (como proteínas) satisfarían las necesidades de nitrógeno de la larva en crecimiento. Con el transcurso del tiempo, por acción de la selección natural, los insectos mejorarían su capacidad de eludir la canavanina y la canalina y, en última instancia, lograron aprovechar su nitrógeno.

Estos éxitos adaptativos fomentarían la dependencia del coleóptero con respecto a *D. megacarpa*. Los beneficios de su asociación con la planta le llevarían a la imposibilidad de poner sus huevos en otras semillas del hábitat. Ante el estrecho espectro alimentario de *C. brasiliensis* y el peligro que comporta la dependencia de un único recurso alimentario y de oviposición, habrá que detenerse en las ventajas potenciales de tal adaptación. El insecto se halla bastante libre de competencia por las semillas y ahorra el gasto que se le exigiría si hubiera de neutralizar una amplia gama de toxinas vegetales. Las larvas en desarrollo tienen asegurado un refugio razonablemente seguro, puesto que la toxicidad de la canavanina debe proteger a las semillas frente a las larvas de otras especies.





8. POSIBLES VIAS DE BIOSINTESIS desde el ácido glutámico (*color*) hasta los aminoácidos marcados con N-15 producidos por la larva de *C. brasiliensis* (*a*). Los aminoácidos que incorporaron nitrógeno pesado se indican mediante rectángulos. Se muestra asimismo la formación de ácido 2-aminobutírico a partir de metionina y homoserina; ambos aminoácidos contenían N-15 (*b*). La síntesis de homoserina por *C. brasiliensis*, el único insecto del que se sabe que puede realizarla, no pudo haberse producido mediante la degradación de canavanina, a través de canalina, puesto que la canavanina no se había marcado.

Un punto final al que nos llevan estos estudios es la notable capacidad de los insectos para adaptarse a un cúmulo de sustancias perjudiciales, incluidos los insecticidas fabricados por el hombre. ¿Acaso los microbios que viven simbióticamente en el intestino del insecto intervienen de una manera decisiva en el procesamiento de compuestos tóxicos como la canavanina? La semilla de *D. megacarpa* muestra un conjunto de simbiontes y de patógenos microbianos que bien pudiera haberlo adquirido *C. brasiliensis*; estos microbios explicarían algunas adaptaciones bioquímicas precisas, si no todas, para la utilización de la canavanina y la canalina. Hemos aislado simbiontes microbianos de las larvas de los escarabajos y descubierto que algunos subsisten con canavanina y canalina (o una sola de ellas) como única fuente de carbono y nitrógeno. Estoy ahora trabajando en la cuestión inmediata: qué parte de la capacidad bioquímica de este depredador de semillas procede de sus simbiontes microbianos.

Nuestras continuadas investigaciones de la relación entre el coleóptero brúquido y su leguminosa patrón nos han permitido esbozar las líneas fundamentales de la estructura bioquímica de la adaptación de un insecto a las toxinas que elabora su planta nutricia. Es una pena que la falta de trabajos bioquímicos similares no extienda la información actual a otras asociaciones. A pesar de todo, lo que hemos encontrado sirve para comprender la respuesta de los insectos a las toxinas naturales y a las medidas químicas de control. Además, estos estudios apoyan la conveniencia de unir en una misma línea de investigación planteamientos ecológicos y bioquímicos.

La erupción de Krakatoa

Comienzan a explicarse las explosiones que hace 100 años arrasaron la mayor parte de esa isla de Indonesia. Los depósitos volcánicos y la cronometría de las ondas atmosféricas y marinas aportan los principales elementos de juicio

Peter Francis y Stephen Self

En la mañana del 27 de agosto de 1883, una serie de explosiones volcánicas a rachas culminó con la sacudida paroxística que acompañó a la destrucción de la mayor parte de Krakatoa, una pequeña isla del estrecho de la Sonda. Los tsunamis (maremotos) desencadenados por la erupción mataron a más de 30.000 personas en las vecinas islas indonesias de Java y Sumatra. El acontecimiento atrajo la atención mundial y las subsiguientes investigaciones científicas despertaron, con importantes aportaciones, el abandonado campo de la vulcanología. Sin embargo, hasta ahora no habían podido explicarse algunos de los principales acontecimientos de la secuencia eruptiva en función de los procesos volcánicos subyacentes. Abordamos aquí las tres cuestiones más debatidas del origen de la erupción: ¿Qué la provocó? ¿Por qué hubo tantas explosiones violentas? ¿Cuál fue la relación entre los devastadores tsunamis y las grandes explosiones?

Cuando Krakatoa entró en erupción, las explosiones se escucharon desde Australia central, Manila, Sri Lanka y hasta en la isla Rodrigues, a más de 5000 kilómetros de distancia, en el océano Indico. En todo el mundo se detectaron ondas atmosféricas de baja frecuencia (demasiado baja para ser audible); en Tokio, a 5863 kilómetros de distancia, los barómetros registraron un aumento de presión de 1,45 milibares. Las olas generadas por la erupción atravesaron el Pacífico y cruzaron el Atlántico: se detectaron con mareógrafos en el golfo de Vizcaya, a 17.000 kilómetros de distancia. El polvo y los gases inyectados en la atmósfera por la erupción provocaron, en los meses subsiguientes, espectaculares puestas de sol en todo el mundo. Las temperaturas medias registradas en el hemisferio Norte durante el mismo período fueron entre 0,5 y 0,8 grados Celsius inferiores a lo normal.

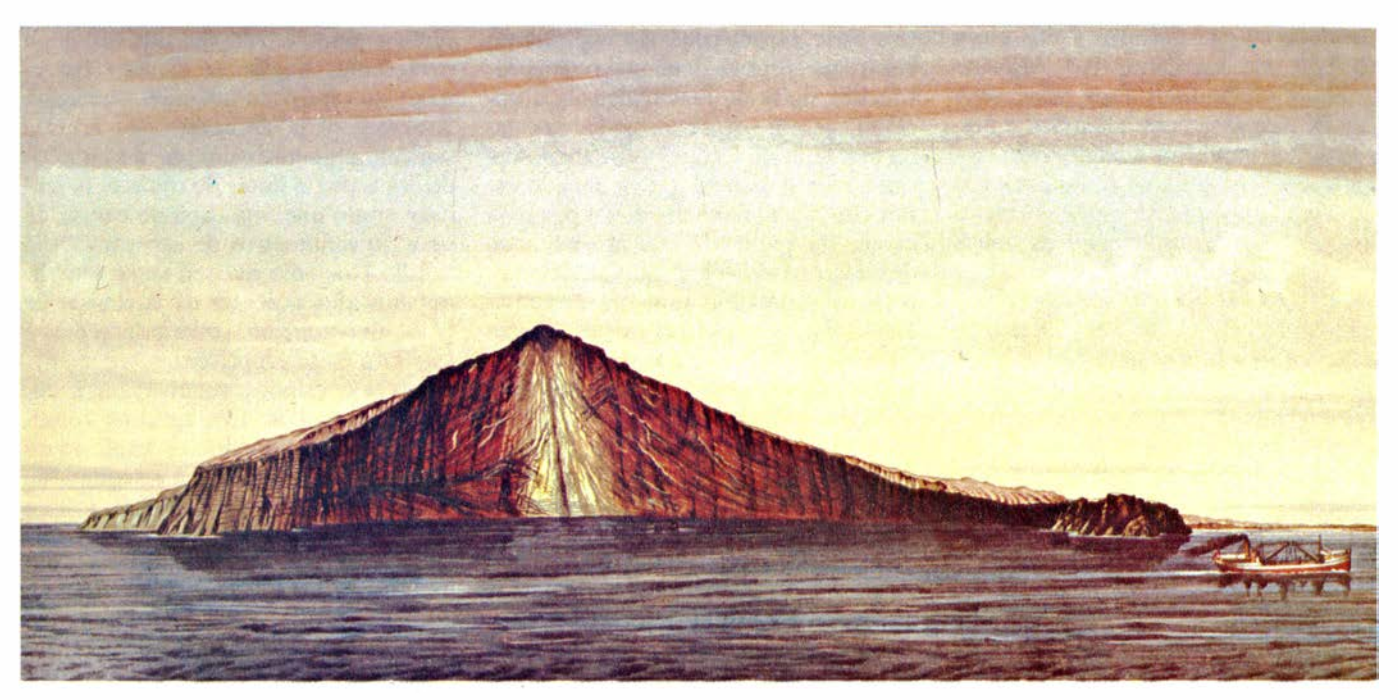
La reputación de Krakatoa como erupción volcánica clásica no se debe sólo a la fecha en que ocurrió; se la ganó también por su violencia. La erupción, una de las primeras en ser objeto de una intensa investigación científica, se produjo en plena era victoriana, cuando la ciencia disfrutaba de un auditorio amplio y entusiasta. La erupción del volcán Tambora, en el archipiélago indonesio de Sumbawa, en 1815, despertó, en comparación, escaso interés, a pesar de que fue mucho más potente. Se estima que Tambora proyectó entre 150 y 180 kilómetros cúbicos de pumita y ceniza; Krakatoa vomitó sólo 20 kilómetros cúbicos. La erupción de Tambora ocasionó la muerte de más de 90.000 personas, directamente o a consecuencia de los tsunamis y del hambre que le siguieron. En aquella época, sin embargo, nadie relacionó el tiempo excepcionalmente frío de Europa y América del Norte del verano de 1816 con la erupción de Tambora del año anterior. (La erupción está necesitada de un estudio serio.)

La Royal Society de Londres y el gobierno holandés, a la sazón administrador colonial de las islas de Indonesia, publicaron extensas memorias sobre la erupción de Krakatoa poco después de producirse. La memoria de la Royal Society recalca los efectos atmosféricos mundiales de la erupción. De sus 494 páginas, 312 se dedicaron a los "insólitos fenómenos ópticos de la atmósfera, 1883-6, incluyendo efectos crepusculares, aparición de coronas, neblinas, soles y lunas coloreados, etcétera". La memoria publicada por el comité holandés, dirigida por el ingeniero de minas y geólogo Rogier D. M. Verbeek, describe los aspectos geológicos de la erupción con mayor detalle. Los miembros del comité visitaron el escenario de la erupción el día 15 de octubre de 1883 y en varias ocasiones más después de esa fecha. Cartografiaron las nuevas islas y los restos de las antiguas y midieron las variaciones de las curvas batimétricas del fondo oceánico. El propio Verbeek recogió muestras de las proyecciones volcánicas, que su equipo y él examinarían mejor al microscopio.

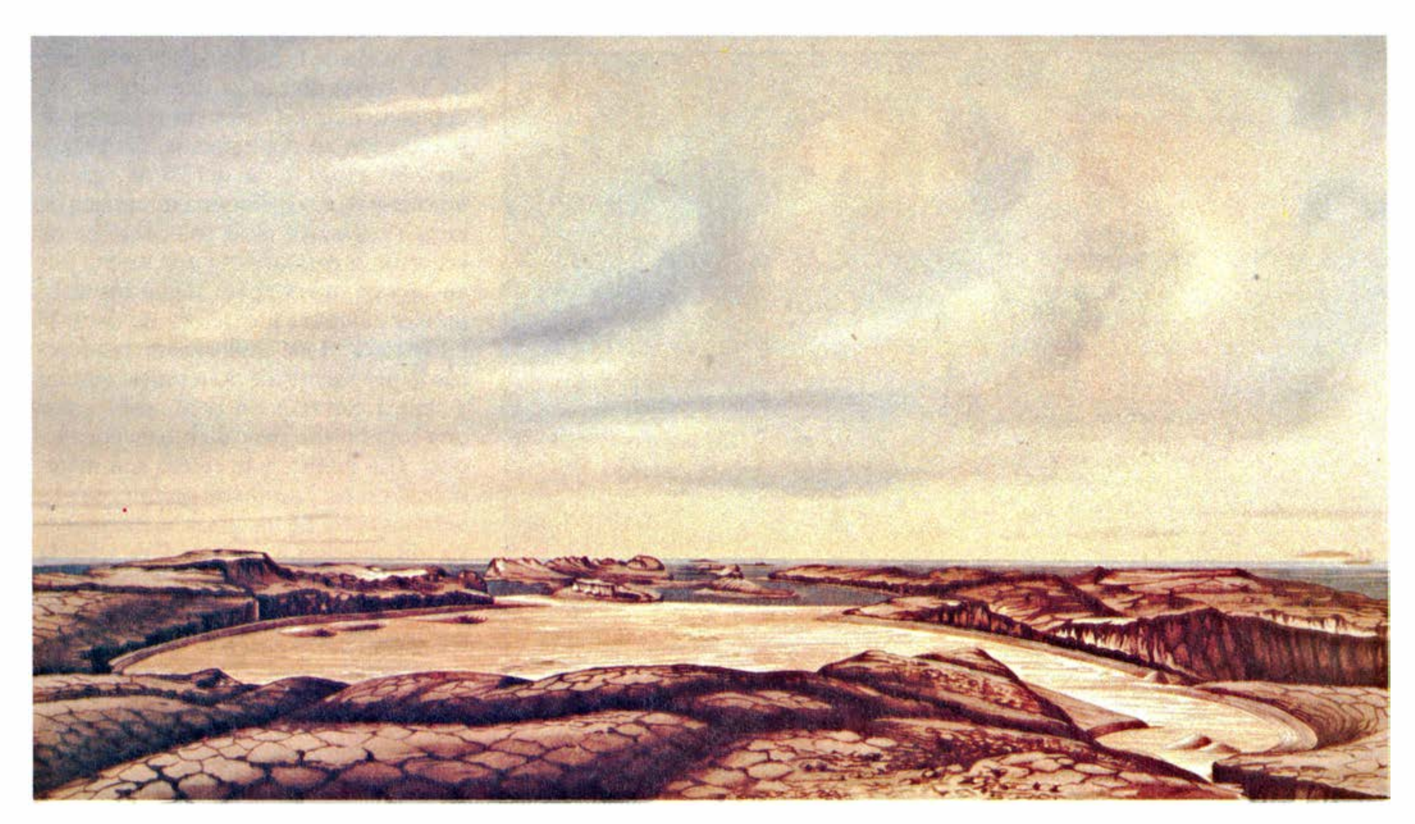
Las conclusiones que Verbeek extrajo de todas aquellas pruebas serían, en su mayor parte, extraordinariamente sagaces. Por ejemplo, basándose en la exploración batimétrica y en la medición del área de lluvia de ceniza hizo una estimación de la cantidad de material arrojado por el volcán tan ajustada que se ha aceptado sin modificación substancial hasta la actualidad. Encontrándose con que, entre las muestras que había recogido, predominaba el material magmático sobre la roca antigua, propuso correctamente que el viejo cono volcánico no había volado por los aires, sino que se había hundido en el mar al agotarse la cámara magmática subyacente.

La historia detallada de la erupción, basada en relatos de los testigos, merece una lectura atenta. Nos centraremos aquí en los sucesos de la secuencia eruptiva que parecen correlacionarse con el emplazamiento de depósitos volcánicos, ya que es en esa correlación donde deben fundarse las reconstrucciones de lo que ocurrió.

La secuencia eruptiva se ha establecido en su mayor parte a partir de los documentos recopilados por los administradores holandeses que vivían muy por encima de la línea de costa en Java y Sumatra, o en ciudades del interior, y a partir de las memorias elaboradas por los oficiales de guardia en los barcos que atravesaban el estrecho de la Sonda; algunos de ellos navegaban cerca de Krakatoa durante la etapa más violenta de la erupción. La descripción de la naturaleza y estratigrafía de los depósitos volcánicos se basa fundamentalmente en el estudio de campo reali-



1. UN CONO VOLCANICO BISECADO es todo lo que quedó de la isla de Krakatoa, situada en el estrecho de Sonda, después de la erupción del 27 de agosto de 1883. La isla estaba formada originariamente por tres conos volcánicos, alineados aproximadamente de noroeste a sureste. Parece que la chimenea principal de las erupciones de 1883 se encontraba entre los dos conos más septentrionales. En las últimas fases de la secuencia eruptiva, dos tercios de la isla se sumergieron en el mar al derrumbarse el techo de la cámara magmática infrayacente. La cara norte del cono meridional, Rakata, que quedó erguida en el borde de la nueva caldera (depresión submarina formada por el desplome), perdió la base de apoyo, y se derrumbó en el mar. Se aprecian claramente la estructura interna del viejo cono volcánico (con rocas blanqueadas por alteración hidrotermal en las inmediaciones de la chimenea central), los diques de alimentación (columnas oscuras), que descienden hasta la cámara magmática, y las capas alternantes de coladas de lava y niveles de ceniza, que constituyen el cono. También se distingue en esta litografía del *Album de Krakatoa* la capa de pumita blanca depositada a ambos flancos del cono.



2. DOS ISLAS NUEVAS, Steers y Calmeyer, aparecieron en el estrecho de la Sonda, al norte de Krakatoa. Se formaron por la exposición, por encima del nivel del mar, de los materiales depositados en el fondo del estrecho, vertidos en sucesivas coladas piroclásticas: nubes de pumita y ceniza pegadas al suelo, arrastradas por gravedad y fluidificadas por gases calientes. Las coladas avanzaron una distancia media de 15 kilómetros desde la chimenea, gran parte del trayecto por debajo del agua o sobre la superficie. El material caliente provocó la vaporización explosiva del agua; algunas de las numerosas explosiones de gran potencia que se registraron ya avanzada la secuencia eruptiva quizá fueran explosiones secundarias de este tipo. Esta cromolitografía del *Album de Krakatoa* muestra un gran cráter de explosión secundaria de Calmeyer. Recuerda mucho a los que se formaron en las coladas piroclásticas del monte Santa Elena que penetraron en el lago Spirit. Los depósitos se erosionaron muy pronto y las dos islas desaparecieron bajo las aguas.

zado por uno de nosotros (Self) y por Michael R. Rampino, del Instituto Goddard de Estudios Espaciales, de la NASA. Estudiamos y tomamos muestras de los depósitos cuando visitamos la isla, en 1979, durante el desarrollo de una amplia investigación de los efectos atmosféricos de las erupciones volcánicas.

Las cartas de navegación de Krakatoa y áreas colindantes del estrecho de la Sonda, levantadas antes de la erupción, muestran que la isla constaba de tres conos volcánicos alineados aproximadamente de noroeste a sureste. El mayor de los conos, Rakata, de 813 metros de altura, se hallaba en el extremo sur de la cadena. Un cono más pequeño, Danán, quedaba en el centro y, el menor, Perbuwatán, en el extremo septentrional. Las coladas de lava que rodeaban Perbuwatán denunciaban actividad en un pasado geológico reciente; allí se registró una erupción de pumita en 1680. Dos pequeñas islas próximas a Krakatoa –Sertung y Rakata Kecil (Rakata Pequeña)– y el extremo sur de la propia Krakatoa probablemente fueran retazos del borde de la caldera sumergida: un gran cráter volcánico formado por derrumbe.

Cuando Perbuwatán despertó con una serie espectacular de explosiones ensordecedoras el 20 de mayo de 1883, tras dos siglos de inactividad, la erupción llegó por sorpresa. Krakatoa estaba deshabitada y sólo la visitaban ocasionalmente pescadores y leñadores. Por ello, cualquier actividad a pequeña escala que pudiese haber precedido su despertar pasó desapercibida. Hubo un período de notable aumento de actividad sísmica en torno al estrecho antes de la erupción, pero a nadie se le ocurrió establecer correlaciones entre esta actividad y Krakatoa.

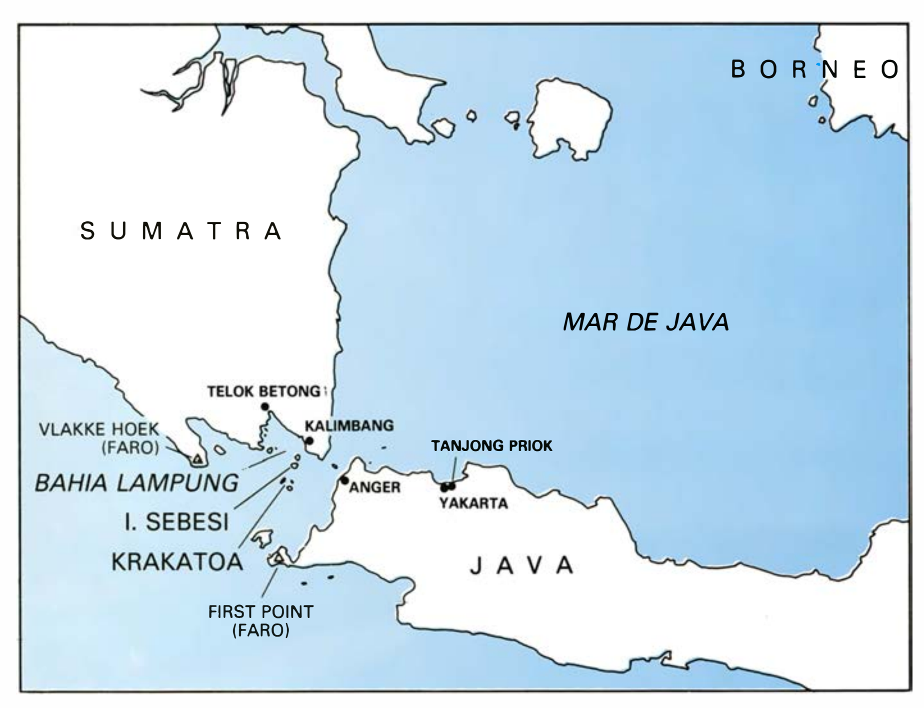
La erupción de mayo del Perbuwatán llegó acompañada de explosiones que se oyeron a más de 150 kilómetros de distancia. A esa lejanía, oscilaciones de presión atmosférica de longitud de onda muy larga tuvieron energía suficiente para parar relojes, sacudir puertas y ventanas y tirar lámparas del techo. Puesto que las ondas de presión eran inaudibles, sus efectos solían atribuirse a terremotos. Aunque se registró alguna actividad sísmica en agosto, durante la fase culminante de la erupción, en este estadio casi toda la energía parece que se transmitió por el aire.

El Perbuwatán continuó su erupción intermitentemente a lo largo de mayo, junio y julio, si bien con una actividad irrelevante. Según el capitán Ferzenaar, un topógrafo holandés que visitó Krakatoa el 11 de agosto, los árboles seguían erguidos, aunque despojados de hojas por la lluvia de cenizas. Ferzenaar anotó que una capa de ceniza de unos 50 centímetros de espesor cubría la isla. Hoy sólo pueden verse muy limitados afloramientos de la ceniza, de grano fino a medio, correspondientes a esta fase de la erupción.

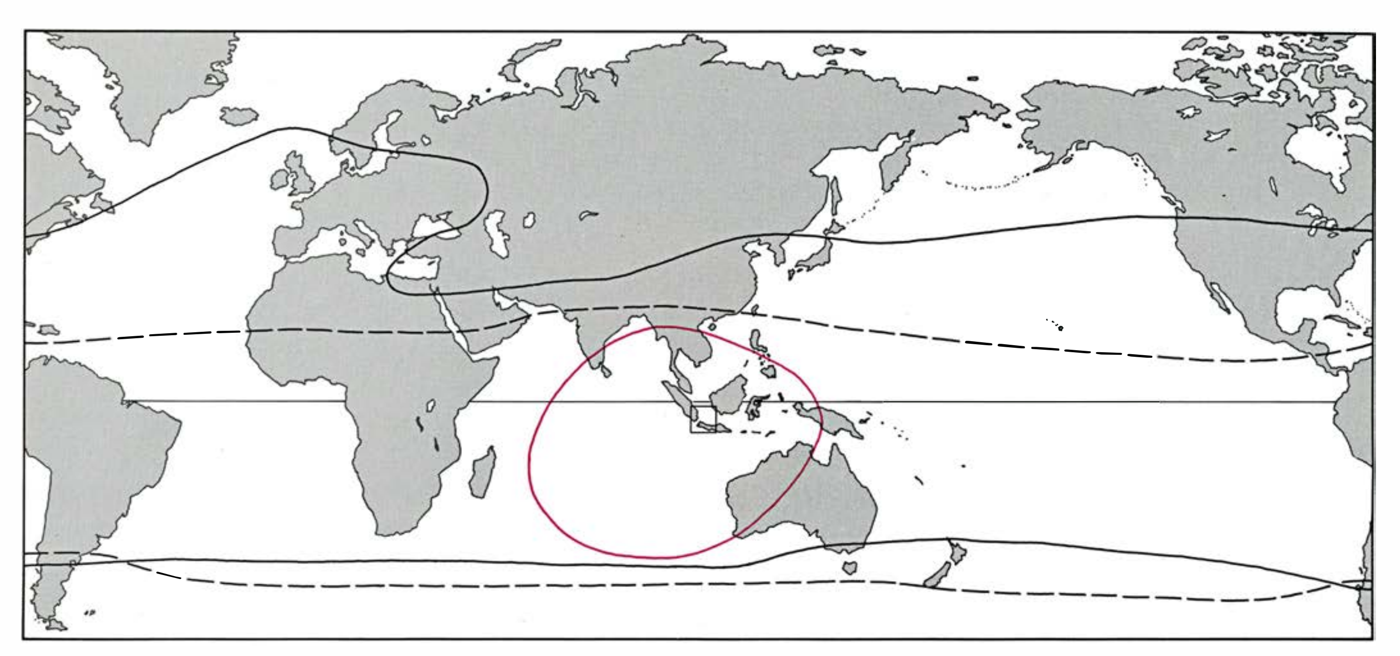
Las explosiones relativamente pequeñas persistieron en agosto y culminaron en las imponentes explosiones del 26 y 27 de ese mes. La actividad fue entonces de magnitud tal que resulta difícil atar cabos para reconstruir exactamente lo que ocurrió; muchos detalles seguramente se han perdido. Nadie que se encontrara cerca del volcán sobrevivió. Las reconstrucciones de la secuencia eruptiva en ese crucial período se basan fundamentalmente en registros instrumentales y en los depósitos volcánicos. Las enormes explosiones del 27 de agosto generaron olas de aire fortísimas, que se grabaron en un manómetro registrador de la fábrica de gas de Yakarta, que conservó así un registro de su duración y amplitud relativa. La duración y magnitud de los tsunamis se registraron en los mareográfos instalados en el estrecho.

En razón de la naturaleza y secuencia de las capas de ceniza depositadas podemos dividir las fases culminantes de la erupción en dos estadios. Empezando a las 13:00 horas del 26 de agosto, una serie de explosiones a intervalos de unos 10 minutos creó una columna de erupción más o menos permanente por encima de la isla, que, según los testimonios, alcanzó una altura de unos 25 kilómetros. Las explosiones produjeron la proyección de materiales: pumita y ceniza elevadas en la atmósfera por una columna de gas caliente en convección. (La pumita y la ceniza son materiales vítreos esponjosos originados por enfriamiento de magma burbujeante. La diferenciación entre ellos atiende fundamentalmente al tamaño; los fragmentos menores de dos milímetros suelen considerarse cenizas.)

Aunque los depósitos de pumita y ceniza procedentes de esta fase de la erupción se acumularon en espesores de más de 20 metros en las islas de Sertug y Rakata Kecil, y los barcos que se encontraban a menos de 20 kilómetros del volcán informaron de una abundante caída de ceniza acompañada de grandes clastos (fragmentos) de pumita que sobrepasaban los 10 centímetros de diá-



3. KRAKATOA ERA UNA PEQUEÑA ISLA DESHABITADA, situada unos 32 kilómetros al oeste de la parte más angosta del estrecho de la Sonda, entre las islas indonesias de Sumatra y Java. Más de 30.000 víctimas se cobró la erupción, que destruyó la mayor parte de la isla. La mayoría las abatieron los tsunamis que barrieron las costas bajas de las islas próximas. No obstante, a lo largo de la costa este de la bahía de Lampong, al sur de Sumatra, se encontraron algunos cuerpos sepultados en ceniza, probablemente de personas que murieron por efectos de coladas piroclásticas, que habían recorrido sobre la superficie del mar más de 40 kilómetros, desde Krakatoa, y que todavía guardaban suficiente calor para abrasar. Los nombres del mapa corresponden a las localidades mencionadas en el gráfico de la secuencia eruptiva que se muestra en la figura 6; los modernos han reemplazado a sus equivalentes del siglo XIX.



4. EFECTOS GLOBALES de la erupción de Krakatoa, representados en un mapamundi. Se ha indicado la extensión afectada por la lluvia de cenizas (*trazo azul*), la extensión en la que se oyeron las explosiones (*trazo rojo*) y la extensión en la que los efectos atmosféricos causados por la ceniza volcánica y los aerosoles inyectados en la alta atmósfera se describieron antes del 22 de septiembre de 1883 (*trazo negro discontinuo*) y hasta finales de noviembre de ese año (*trazo negro continuo*). (En recuadro negro se señala el territorio que cubre la figura 3.) La ceniza de la erupción cayó sobre una superficie de 700.000 kilómetros cuadrados, la mayor extensión de lluvia de cenizas jamás producida por una erupción volcánica en tiempos históricos. La explosión mayor originó densas nubes de ceniza fina sobre el estrecho de la Sonda, que probablemente atenuaron el ruido con gran eficacia. Quizá por ello la gente que vivía en los litorales de Java y Sumatra, que pronto serían víctimas de los tsunamis originados por la erupción, no oyeron las explosiones, que sí percibieron los habitantes de lugares tan distantes como Australia central. Un estudio llevado a cabo por el físico británico Rollo Russell sobre la trayectoria y velocidad de la nube alta, basado en informes de efectos atmosféricos, proporcionó la primera prueba para los modelos de circulación estratosférica.

metro, la densa lluvia de ceniza tuvo una extensión limitada. En Sumatra y en la parte occidental de Java la lluvia de ceniza fue escasa.

A las 5:30 horas del día 27 de agosto, el carácter de la erupción cambió espectacularmente. A lo largo del día hubo muchas explosiones de enorme potencia. La mayor explosión, la que se oyó hasta en la isla de Rodrigues, se produjo a las 9:58 horas; estuvo asociada al mayor de los tsunamis, cuya cresta se estima en unos 40 metros y provocó la mayoría de las bajas sufridas en las costas cercanas.

Durante este período, las explosiones se caracterizaron por la formación de coladas piroclásticas, así como una columna permanente de material proyectado en el aire. Las explosiones del 27 de agosto, más que continuas, parece que fueron paroxísticas. Cada sacudida proyectaba gran cantidad de pumita y ceniza por los aires, a veces a alturas de hasta cinco kilómetros. La consistencia del material, sin embargo, era demasiado densa y pesada para permanecer suspendida en el aire; en su mayor parte caía inmediatamente al suelo, donde formaba nubes bajas incandescentes que, arrastradas por la gravedad y fluidificadas por gases calientes, se desplazaban por la isla y el mar. Parte de la ceniza se elevó aún más por las corrientes de convección creadas en la atmósfera por calentamiento local. Las nubes de ceniza resultantes debieron alcanzar una altura de unos 40 kilómetros. La ceniza de esas nubes cayó sobre una extensa área; las costas cercanas se sumieron en la oscuridad y se registró lluvia de cenizas en lugares tan alejados como las islas Cocos, a 1850 kilómetros de distancia de Krakatoa.

Las coladas piroclásticas dejan un depósito característico llamado ignimbrita; tales depósitos constituyen la mayor parte del material arrojado por Krakatoa. Dado que el material de las coladas piroclásticas estaba fluidizado (presentaba densidad y viscosidad bajas) por gases calientes, hacían aquellas un uso extraordinariamente eficaz de la energía cinética adquirida al caer de lo alto de la columna de erupción, y recorrían largas distancias horizontales. Se sedimentaron depósitos de ignimbritas de hasta 40 metros de espesor, extendiéndose hasta 15 kilómetros de distancia del emisor primario, que, según se cree, estaba entre Danán y Perbuwatán.

Parece que las coladas piroclásticas se propagaron preferentemente hacia el norte y noroeste, cubriendo las islas y el fondo del mar circundante con un manto de ignimbrita. La distribución probablemente se debió al elevado cono de Rakata, el cual empujaba hacia el norte el material que caía de la columna de erupción. Los relatos de incendios causados por cenizas calientes procedentes del área de alrededor de Kalimbang, al sur de Sumatra, prueban que algunas de las coladas avanzaron hasta 40 kilómetros al nordeste. Está claro que los daños se produjeron por flujos horizontales, y no los causó la caída vertical de ceniza: en un caso, los supervivientes describieron que los gases calientes soplaban a través de las tablas del suelo de la casa.

El 28 de agosto, el barco holandés *Gouverneur-Generaal Loudon* zarpó de Telok-Betong, en Sumatra, rumbo a Yakarta, en Java, a través del estrecho de la Sonda, al norte de Krakatoa. El estrecho apareció de pronto demasiado somero en algunos puntos para permitir la navegación; en otros, islas flotantes de pumita lo bloqueaban. El barco se vio obligado a desviarse considerablemente de su ruta acostumbrada hacia el este, navegando hacia el oeste y finalmente pasando por el sur de Krakatoa. La ruta del barco siguió aproximadamente el borde exterior de la ignimbrita que se había depositado en el fondo del mar.

El mapa hidrográfico detallado del

RAKATA
GRABEN
RAKATA KECIL
DANAN
PERBUWATAN
CALDERA ANTES DE 1883
SERTUNG

RAKATA
CALDERA ANTES DE 1883
CALMEYER
CALDERA DE 1883
IGNIMBRITA
STEERS
N

equipo de Versbeek, levantado a mediados de octubre, registra los cambios topográficos producidos por la erupción. Los dos tercios septentrionales de la isla de Krakatoa habían desaparecido. Las líneas de costa de Sertung y Rakata Kecil se habían ensanchado hasta tres kilómetros con los depósitos de pumita y ceniza. Además, la parte sur, no sumergida, de Krakatoa apareció rodeada de depósitos de ignimbrita blanca. La mayoría del material volcánico, sin embargo, fue a parar al mar. Zonas del estrecho de la Sonda, que antes alcanzaban de 20 a 60 metros de profundidad, se rellenaron de ignimbrita. En el estrecho, hacia el norte de Krakatoa, la ignimbrita emergida por encima del nivel del mar formó dos islas, más tarde llamadas Steers y Calmeyer.

En la vieja caldera prehistórica se formó una nueva, que en su borde meridional se hallaba a 290 metros de profundidad y que colmataba parcialmente la antigua. La distribución de los depósitos de flujo piroclástico en los lados norte y este de Rakata sugiere que se depositaron antes de formarse la caldera, lo que viene a indicar que la cámara magmática se derrumbó en una fase tardía de la secuencia eruptiva. La mayor parte de la isla se hundió en el mar cuando cedió el techo de la cámara, pero el cono de Rakata quedó, aparentemente, colgado del borde sur de la caldera, con su cara norte virtualmente sin apoyo. El flanco del volcán se deslizó entonces en el mar, dejando tras de sí un cono volcánico espectacularmente bisecado, todo lo que queda de la isla de Krakatoa.

El escaso progreso conseguido hasta hace poco en la comprensión de lo que ocurrió en Krakatoa se explica en parte por la poca atención prestada a las características físicas de los depósitos volcánicos, en especial a los dejados por las coladas piroclásticas. Verbeek estudió la petrología de los depósitos y C. E. Stehn (quien visitó el área en 1927, cuando la erupción de un nuevo volcán, aún submarino, en el interior de la caldera rebasó la superficie) estudió su estratigrafía. Howel Williams, de la Universidad de California en Berkeley, realizó una importante aportación en 1941, al indicar que la mayor parte de la pumita expuesta se había emplazado por coladas piroclásticas, y no por la caída de material lanzado a la atmósfera. Sólo en los últimos 15 años, sin embargo, los vulcanólogos han empezado a descifrar el significado de ciertas características físicas de los depósitos volcánicos: tamaño de grano y estructura interna, entre ellas. Estos rasgos de los depósitos de Krakatoa se estudiaron por primera vez cuando uno de nosotros (Self) y Rampino visitaron la isla en 1979.

En cierto modo no debe sorprender que esos aspectos de los depósitos no se examinaran antes. La mayoría de la ignimbrita se depositó en el mar, y las islas de ignimbrita, Steers y Calmeyer, se erosionaron rápidamente y desaparecieron de la superficie. Los depósitos caídos en las islas viejas también se erosionaron enseguida. Verbeek observó, dos meses después de la erupción, que en los depósitos se habían excavado barrancos de laderas escarpadas de 40 metros de profundidad. La topografía, sumamente irregular, quedó envuelta en una exuberante y pronto impenetrable vegetación. Hoy, los depósitos sólo pueden muestrearse desde embarcaciones, al pie de los escarpes costeros. La mayoría de las cuestiones sobre la erupción todavía giran en torno a las características físicas de los depósitos.

El estudio de los depósitos proporciona la base de nuestras respuestas a la primera cuestión: ¿qué provocó la erupción? Las principales erupciones volcánicas han tenido su causa en la súbita descompresión de magma, que está saturado o sobresaturado en gases, como dióxido de carbono y vapor de agua. Los volátiles disueltos pueden desprenderse por dos vías. La presión de la cámara magmática crece gradualmente cuantos más volátiles abandonan la solución; cuando la presión supera la resistencia de las rocas suprayacentes, el magma fuerza una salida a la superficie. Segunda vía: a través de un proceso tectónico, así el corrimiento de tierras forzado por un terremoto o el movimiento a lo largo de fallas localizadas por encima de la cámara magmática, que crea una abertura que provoca la descompresión instantánea. No hay pruebas de que ningún proceso tectónico desencadenara la erupción del Krakatoa, por lo que la explicación debe buscarse en procesos desarrollados dentro de la propia cámara magmática.

Verbeek sugirió que las explosiones finales empezaron cuando penetró agua del mar en la cámara magmática y reaccionó violentamente con el magma caliente. Esta hipótesis se aceptó ampliamente en cierta época, y se generalizó hasta incluir las primeras explosiones. Las erupciones volcánicas provocadas por la interacción violenta del agua con el magma, llamadas erupciones freatomagmáticas, pueden tener lugar, en efecto, cuando el magma calienta y fractura la roca que encierra un manto de agua subterránea, o cuando el agua de mar consigue penetrar en la cámara magmática. La mezcla de agua subterránea o del mar con el magma, altamente explosiva, suele originar depósitos característicos de ceniza de grano muy fino y de gran dispersión.

Cabe esperar que la intrusión de grandes cantidades de agua de mar enfríe el magma, de modo que el material expulsado, especialmente la ceniza más fina, se refrigere antes en una erupción freatomagmática que en otro tipo de erupciones. Los datos de campo recogidos en Krakatoa son algo ambiguos en este aspecto. Por un lado, los primeros depósitos caídos del aire en Rakata Kecil, sólo a 2,3 kilómetros de la supuesta chimenea, estaban tan calientes que los fragmentos de pumita vítrea poseían blandura suficiente para soldarse unos con otros tras su deposición. La naturaleza de estos depósitos así acumulados hace improbable que el material proyectado se hubiese enfriado por contacto con agua.

Por otro lado, George P. L. Walker, de la Universidad de Manoa en Hawai, ha señalado que la naturaleza de algunos de los depósitos de ignimbrita revelan que el magma se enfrió con agua. Los núcleos de los grandes clastos de pumita son más porosos que la periferia. Además, los clastos tienen una costra exterior que podría haberse formado por enfriamiento súbito. Estos datos sugieren que los clastos estuvieron inicialmente más calientes que la matriz de ceniza que los rodeaba, y que se enfriaron de fuera adentro. Pero cabe también que el agua de mar enfriase la

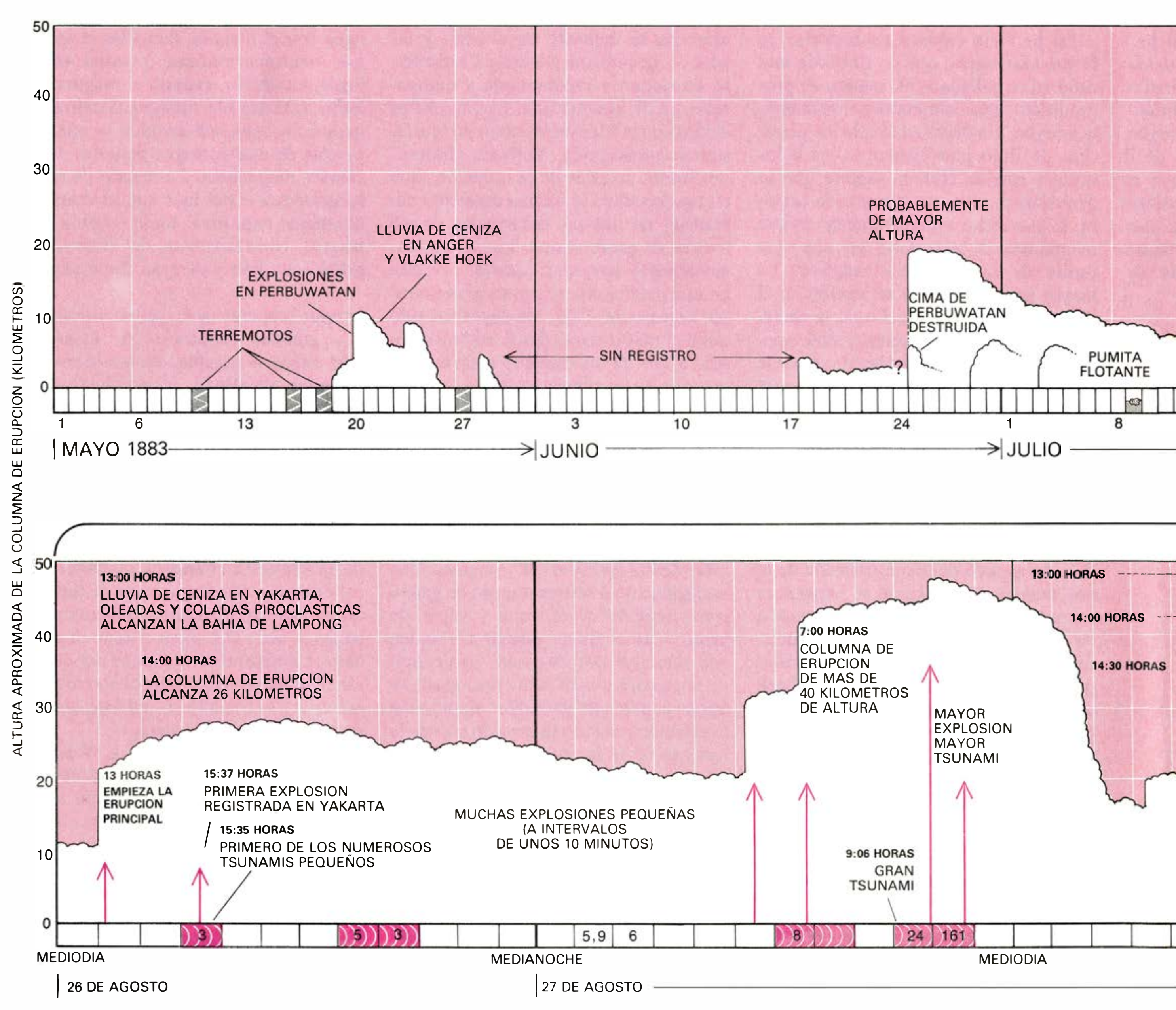
5. DIBUJOS DE KRAKATOA, antes y después de la erupción, reconstruidos a partir de mapas hidrográficos contemporáneos. La isla estaba formada por tres conos volcánicos: Rakata, Danán y Perbuwatán. Las islas de Sertung y Rakata Kecil, y tal vez el propio extremo sur de Krakatoa, eran restos del borde de una caldera prehistórica. Dos depresiones submarinas (grabens) próximas a las islas indican que en esa zona la corteza estuvo sometida a esfuerzo en extensión, lo que quizá menguara la corteza suficientemente para alojar la cámara magmática que produjo la erupción de 1883. La mayor parte del magma se proyectó en forma de coladas piroclásticas, que depositaron en el fondo del mar una capa de ignimbrita de hasta 40 metros de espesor. Las coladas se desplazaron preferentemente hacia el norte y el nordeste, porque el alto cono de Rakata actuaría de barrera por el sur. Las dos nuevas islas, Steers y Calmeyer, correspondían a las zonas donde la ignimbrita asomaba por encima del nivel del mar. Cuando la erupción agotó la reserva de magma, la cámara se derrumbó formando una nueva caldera, en la que se hundió la mayor parte de la isla, que desapareció de la superficie. La caldera de 1883 debió de ser alargada, puesto que el derrumbe siguió las líneas de falla de los grabens. (Los dibujos no respetan la escala.)

ignimbrita al salir por la chimenea o al depositarse. (La mayor parte de depósitos visibles en la actualidad están a nivel del mar.)

Las pruebas geológicas, por consiguiente, no indican con claridad que hubiese contacto directo explosivo entre el agua del mar y el magma caliente en la cámara magmática o en el conducto de Krakatoa. No obstante, puesto que la chimenea se hallaba próxima al nivel del mar, parece probable que hubiera explosiones freáticas de pequeña entidad, tal vez ya avanzada la erupción, quizás a partir del atardecer del 26 de agosto; mas no debieron producir contacto directo entre el agua y el magma. Tales explosiones debilitarían el techo de la cámara magmática principal ocasionando una descompresión súbita y la emisión de grandes volúmenes de magma vesicular. Nosotros proponemos, sin embargo, que la interacción del agua y el magma fue una causa adicional, no el motivo principal de la erupción.

La naturaleza de los propios fragmentos de pumita apunta hacia otro mecanismo desencadenante de la erupción más verosímil. El vidrio esponjoso de la pumita incluye un escaso número de cristales minerales, que empezaron a formarse en la cámara magmática antes de la erupción. (Cuando el magma se enfría deprisa, se forma vidrio; el enfriamiento lento favorece la formación de cristales.) La composición de la pumita y de algunas inclusiones cristalinas depende de la composición del magma del que surgieron. La composición de los magmas varía desde basalto, material oscuro relativamente pobre en sílice, hasta riolita, material gris claro rico en sílice, pasando por los intermedios de andesita y dacita. Los magmas de coloraciones blanquecinas se llaman silíceos; forman minerales con alto contenido en sílice, como cuarzo y feldespatos. Los magmas oscuros se denominan máficos; forman minerales ricos en magnesio y en hierro, como el piroxeno.

La mayoría de los depósitos de Kra-

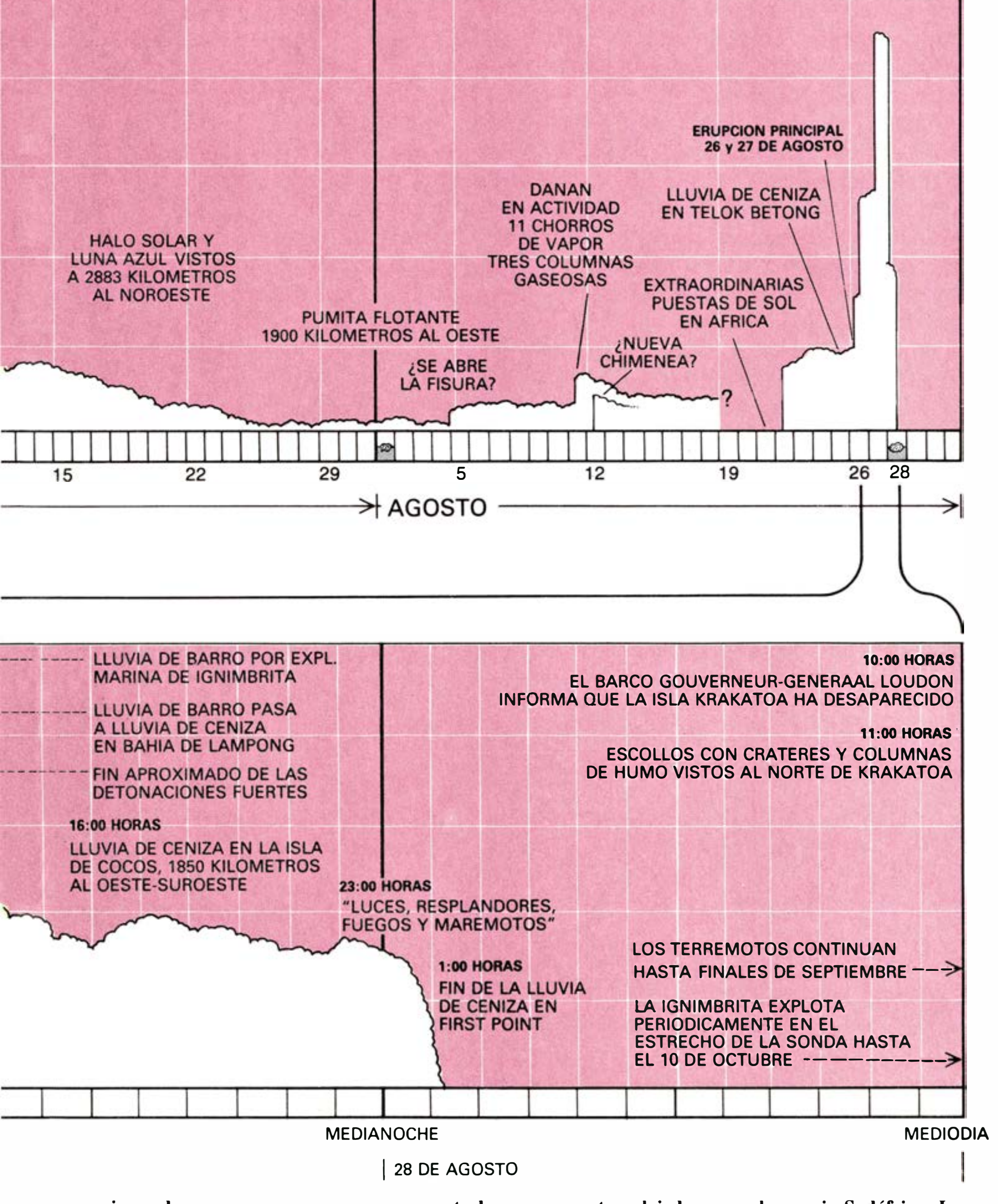


6. SECUENCIA ERUPTIVA DE KRAKATOA, reconstruida a partir de varias fuentes: diarios y cuadernos de notas de los administradores holandeses y otros habitantes de islas próximas, apuntes tomados por oficiales de guardia en barcos que navegaban por el estrecho de la Sonda, acontecimientos registrados por un manómetro de una fábrica de gas de Yakarta y por mareógrafos emplazados a lo largo del estrecho de la Sonda y, ocasionalmente, observa-

katoa están formados por pumita dacítica pálida que contiene pocos cristales minerales. Pero los hay muy distintos: presentan vetas o bandas de vidrio oscuro intercaladas entre el vidrio claro, y algunos clastos son completamente oscuros. La pumita más oscura también incluye cristales de composición distinta de la propia de los de la pumita clara. Tal mezcla de pumitas no es rara en los depósitos piroclásticos. Se ha sugerido que puede formarse cuando su magma basáltico caliente hace intrusión en un magma claro riolítico o dacítico y ambos se mezclan.

La intrusión de nuevo magma basáltico caliente en la base de la cámara de magma dacítico puede causar un violento movimiento convectivo en la cámara. Los magmas silíceos tienden a ser menos densos que la mayoría de los magmas máficos, en gran medida por su diferente composición. Se cree que la mayoría de las cámaras magmáticas dacíticas están estratificadas de acuerdo con su composición, con el magma más silíceo en el techo y los materiales más máficos, algo más densos, debajo. Tal sistema es estable. Stephen Sparks y Herbert Huppert, de la Universidad de Cambridge, han propuesto, sin embargo, que si en la cámara hace intrusión magma basáltico caliente puede iniciarse un movimiento convectivo por sobrecalentamiento del nivel adyacente; éste se hace así menos denso que el magma silíceo que tiene encima. Una segunda posibilidad es que el magma basáltico caliente se mantenga en el fondo de la cámara y la convección violenta se produzca al enfriarse. Entonces, se desprenden los volátiles de la solución y los cristales sedimentan, disminuyendo la densidad del líquido hasta que se hace más baja que la de los niveles superiores.

¿Por qué el movimiento convectivo provoca la explosión? En general, la erupción de magma basáltico tiende a ser menos explosiva que la de la mayoría de magmas silíceos, porque es menos viscoso y las burbujas de gas originadas por los volátiles salen de la solución con facilidad, filtrándose a través de la propia masa o exhalándose; en cambio, las burbujas formadas en magmas silíceos, más viscosos, quedan bloqueadas en su interior y provocan su fragmentación explosiva cuando la presión interna de aquéllas excede su resistencia. En una impresionante serie de experimentos de laboratorio, J. Stewart Turner, de la Facultad de Investigaciones Geológicas de la Universidad Nacional de Australia, y Sparks y Huppert han demostrado la importancia del contenido en volátiles del nuevo magma en el proceso de mezcla, o contaminación, de la cámara magmática. El nuevo magma basáltico puede ser rico en volátiles, mantenidos en disolución por la presión correspondiente a la profundidad a la que se formó dicho magma. El movimiento convectivo transporta el material rico en volátiles hasta niveles menos profundos de la cámara magmática, donde la presión ambiental es menor. Conforme sube el magma, los volátiles escapan rápidamente de la solución, aumenta la presión en la cámara magmática y se posibilita la erupción explosiva.



ciones de personas, que a veces se encontraban en zonas tan alejadas como la propia Sudáfrica. Los números de las casillas de color expresan la magnitud relativa de los tsunamis. La de las explosiones se expresa por la altura de las flechas. Sólo se han representado algunas de las explosiones principales.

Una pieza importante de la documentación histórica en que se basa la hipótesis según la cual la erupción de Krakatoa debióse a contaminación magmática salió a la luz durante la preparación de este artículo. El 27 de mayo de 1883, poco después de la primera de las explosiones del Perbuwatán durante ese mes, un grupo de por lo menos 86 personas, entre ellas el ingeniero de minas holandés J. Schuurman, visitó Krakatoa. Schuurman escribió una memoria detallada de esa visita; señalaba que 60 centímetros de "ceniza" oscura cubrían 30 centímetros de "pumita" de color claro. También recogió muestras del material proyectado, que posteriormente se analizaron. Las

muestras de pumita eran de composición similar a la pumita dacítica arrojada por las explosiones de agosto. La ceniza gris, en cambio, ofrecía una composición basáltica. Así pues, parece probable que en la primera fase de la erupción, poco después de la emisión inicial de magma dacítico, se arrojara una pequeña cantidad de magma basáltico. Ello indica que en mayo subyacían al volcán magmas de composiciones notablemente diferentes, lo que contribuye a confirmar la hipótesis de que las explosiones de agosto se originaron por contaminación magmática resultante de la intrusión de nuevo magma basáltico.

Aunque es imposible probar que fue esa contaminación magmática la que provocó la erupción de Krakatoa, las experiencias de laboratorio de Sparks y Huppert demuestran de manera convincente que el mecanismo es plausible. Eso mismo se ha propuesto para explicar otras erupciones importantes, como la del Askja, en 1875, en Islandia, y la del Santa María, en 1902, en Guatemala, localidades donde se han encontrado también mezclas semejantes de pumitas. Sin embargo, la hipótesis de la contaminación magmática deja un problema sin resolver. Aunque se suponga que la contaminación magmática inició la erupción, no está claro por qué algunas de las explosiones de la secuencia alcanzaron esa violencia.

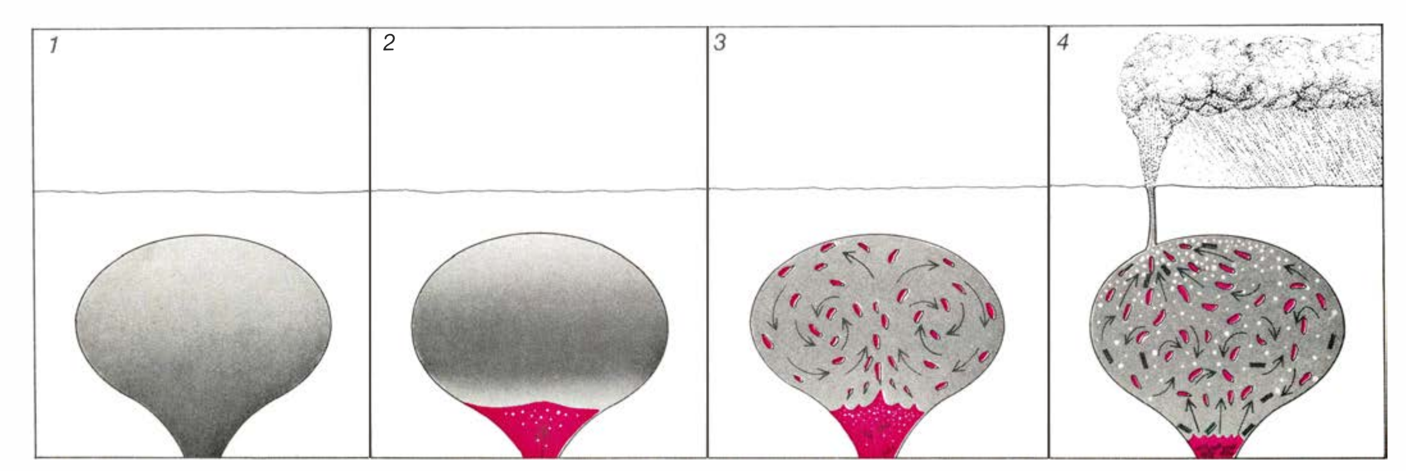
Vayamos con la segunda cuestión planteada: ¿Por qué la erupción de Krakatoa se vio acompañada de múltiples explosiones? Podríamos hallar una primera razón en considerar, algunas de ellas, explosiones secundarias ocasionadas por la penetración de coladas piroclásticas en el mar. No conocemos bien qué ocurre cuando un gran volumen de colada piroclástica entra en contacto con el agua del mar; son contados los ejemplos hasta ahora descritos. ¿Desplaza la colada el agua y se mueve por el fondo del mar, o avanza sobre la superficie del agua? Parece ser que se dan ambas cosas, aunque lo que tenga exactamente lugar dependerá, a buen seguro, de las características de la colada.

Al entrar la colada en el mar, la parte inferior, densa y formada principalmente por pumita, material vítreo de grano fino y fragmentos de rocas no volcánicas más antiguas, debe sumergirse en el agua. La parte superior, más difusa y turbulenta, constituida fundamentalmente por pumita y ceniza fluidificadas por gases calientes, quizás adquiera transitoriamente una densidad global menor que la del agua del mar y se desplace sobre la superficie. Conforme la turbulencia se torna menos eficaz, los clastos mayores se desprenden y el flujo termina por quedarse sin material sólido. Sólo permanece una masa turbulenta de gas y partículas finas; tales flujos, aún calientes, probablemente alcanzaron Kalimbang, en la bahía de Lampong.

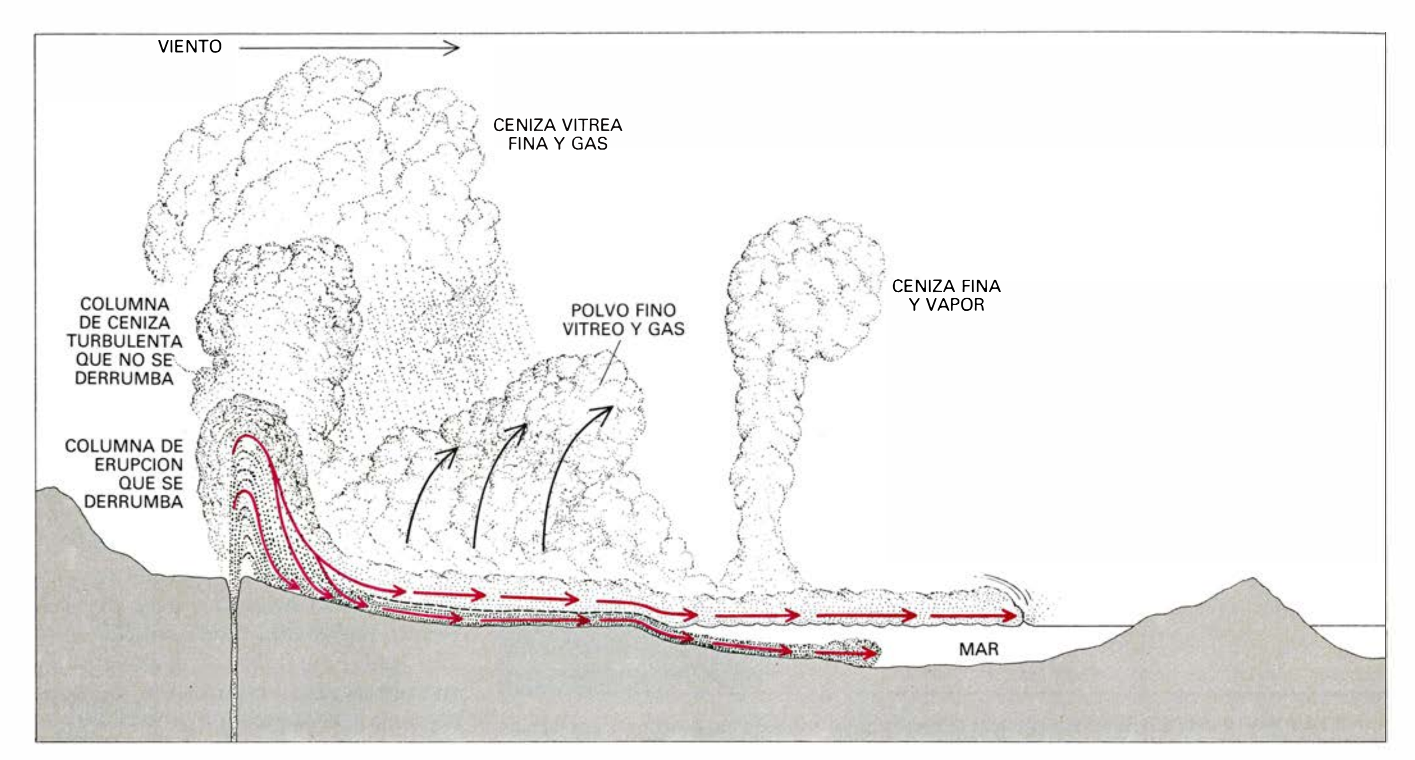
Una colada piroclástica depositada hace 6000 años por la erupción del volcán Koya, en la isla japonesa de Kyushu, aporta pruebas de que la interacción de la ignimbrita caliente con el agua del mar puede ser, en algunas circunstancias, tan explosiva como imaginamos intuitivamente. Walker, basándose en el estudio de estos depósitos, explica que las explosiones ocasionadas por la penetración de coladas en el mar producen lluvias de cenizas sobre muchos miles de kilómetros cuadrados. El carácter distintivo de estos depósitos de ceniza es que su espesor y tamaño de grano varían poco con la distancia del origen, distribución que sugiere que la ceniza se originó en explosiones de extraordinaria fuerza.

Los depósitos de Rotoehu arrojados hace 50.000 años por el núcleo volcánico de Okataina, en Nueva Zelanda, demuestran que esas capas de ceniza pueden presentar escasa cristalinidad. Walker sostiene que los cristales se separan de los fragmentos de su pumita madre por un proceso natural de selección. Mientras la colada corre hacia el mar, los cristales más densos se sumergen y se levanta el polvo de pumita. El aventado de los cristales de la ceniza acompaña las explosiones que se producen cuando la colada penetra en el mar, a la vez que potentes torbellinos tamizan las nubes de ceniza producidas por las explosiones y arrastran el polvo ardiente.

Con este distanciamiento en el tiempo, resulta difícil, por supuesto, determinar exactamente qué ocurrió cuando las coladas piroclásticas penetraron en el mar. Las ilustraciones contemporáneas, sin embargo, muestran grandes



7. UNA MEZCLA VIOLENTA DE MAGMAS de distinta composición pudo provocar la erupción volcánica de Krakatoa. La composición de los magmas abarca desde basaltos pobres en sílice hasta riolita, rica en sílice, pasando por los términos intermedios de dacita y andesita. Los magmas menos silíceos tienden a ser los más densos. La mezcla de magmas dacíticos de la cámara de Krakatoa debió estar originariamente estratificada de acuerdo con su densidad, lo que le conferiría estabilidad (*1*). Probablemente hizo intrusión por la base de la cámara nuevo magma basáltico, más caliente que el que ya había en la cámara y también más rico en sustancias volátiles, y que al principio debió permanecer en el fondo de la cámara (*2*). El magma caliente sobrecalentaría el nivel del magma dacítico situado inmediatamente encima, haciéndolo menos denso que los niveles que lo rodeaban. Al enfriarse el magma basáltico, empezarían a formarse cristales y se desprenderían volátiles de la solución; el magma basáltico sería entonces menos denso que los niveles inferiores de magma dacítico. Cualquiera de estos procesos, o ambos a la vez, pudieron ocasionar la violenta movilización convectiva del material de la cámara (*3*). Cuando el magma basáltico, rico en volátiles, ascendió a niveles de presión ambiente reducida, los volátiles disueltos abandonaron rápidamente la solución. La presión de la cámara magmática aumentaría entonces hasta el punto que el magma se abrió un sendero hacia la superficie e hizo erupción.



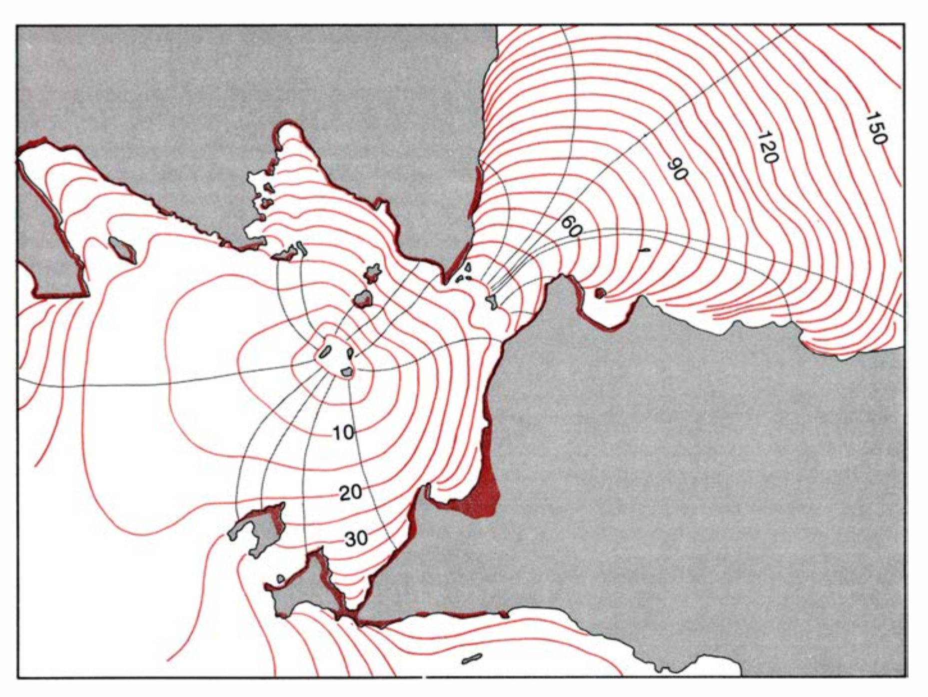
8. LAS COLADAS PIROCLASTICAS son masas de material volcánico incandescente fluidificado por gases calientes (presentan, por tanto, una densidad y viscosidad más bajas que las de una acumulación de material sólido). Conforme la colada se desplaza, se diferencia en su interior una zona baja, más densa, en la que el flujo es laminar, y una zona superior, ligera, en la que el flujo es turbulento. Cuando la colada llega al mar, los materiales más densos se hunden en el agua, pero la parte superior de la colada presenta temporalmente menor densidad que el agua, y por tanto se desplaza sobre la superficie. Al disiparse la fluidización, el material que queda se hace más denso y en su mayor parte penetra en el mar. El proceso continúa hasta que sólo queda una nube turbulenta de ceniza y vapor. El material caliente puede ocasionar explosiones secundarias cuando se sumerge en el agua. Sostienen algunos autores que el proceso de selección que tiene lugar en la propia colada, y el aventado de la ceniza por las explosiones secundarias, producen una ceniza vítrea fina que se distribuye sobre una gran área con las explosiones secundarias. R. Verbeek descubrió depósitos así cerca de Krakatoa.

cráteres circulares en Calmeyer, una de las islas originadas por el depósito de las ignimbritas. Los cráteres se parecen extraordinariamente a los producidos por las explosiones secundarias que ocurrieron cuando las coladas piroclásticas arrojadas por el monte Santa Elena, en mayo de 1980, penetraron en el lago Spirit. Por otra parte, Verbeek observó que los niveles más altos de ceniza de la isla de Sebesi eran de grano extraordinariamente fino y estaban constituidos casi exclusivamente de vidrio. Esta ceniza fina y escasamente cristalina bien pudo provenir de la deposición tras las fuertes explosiones secundarias. Por todo ello proponemos que, aunque las gigantescas sacudidas de Krakatoa ocurridas a las 9:58 de la mañana del día 27 de agosto probablemente respondieran a la expulsión del magma que generó coladas piroclásticas, al menos algunas de las explosiones más fuertes de la secuencia eruptiva siguieron a la penetración de coladas piroclásticas en el mar.

Los mayores estragos los produjeron los grandes tsunamis desencadenados por la erupción. A pesar de la intensa investigación de que han sido objeto, continúan sin determinarse la causa o causas que los provocan. Los tsunamis suelen producirse por un movimiento vertical súbito del fondo del mar, forzado, corrientemente, por un terremoto. En los años 1960 se argumentó que los tsunamis de Krakatoa fueron ocasionados por el mecanismo similar de formación de la caldera. En ese punto el debate se centra en averiguar cuándo y con qué rapidez se derrumbó el techo de la cámara magmática, así como en determinar si los tsunamis estuvieron precedidos por un descenso del nivel del mar, como podría esperarse de la precipitación del agua a llenar la caldera recién formada. No se ha resuelto la datación exacta del derrumbe de la caldera, pero parece que se produjo en una fase bastante avanzada de la secuencia, después incluso de algunos de los grandes tsunamis. Los datos relativos a la segunda cuestión también son algo confusos. En algunos puntos, localizados a lo largo del estrecho de la Sonda, la llegada de la mayoría de los tsunamis se caracterizó por una subida del nivel del mar. Sin embargo, aparentemente se registraron también algunas retiradas, aunque la mayoría de los mareógrafos se hallaban muy lejos de Krakatoa para que una retirada inicial, pequeña, quedara registrada.

Se aducen tres causas posibles de los tsunamis. Para algunos investigadores, las erupciones submarinas, al igual que las cargas de profundidad, podrían crear una cúpula de agua que, al propagarse hacia fuera, provocarían el tsunami. Verbeek propuso en 1884 que el mayor de los tsunamis se había generado por el desprendimiento de la mitad norte del cono volcánico de Rakata hacia el interior de la caldera recién formada. A Verbeek también debemos una tercera hipótesis: el súbito desplazamiento de agua por las proyecciones volcánicas "llovidas" sobre el mar podrían haber iniciado el movimiento de las olas. Dado que fueron muchos kilómetros cúbicos de material los que penetraron en el mar en forma de coladas piroclásticas, esta posibilidad nos parece la más verosímil.

La correlación del tiempo de llegada de las olas a distintos puntos situados a lo largo del estrecho de la Sonda con los acontecimientos de la secuencia eruptiva es crítica a la hora de inclinarse por el origen de las olas. El establecimiento de la cronología se ha visto complicado por dos problemas. Primero, no todos los tsunamis eran verdaderos tsunamis, es decir, ondas que se habían propagado únicamente por el

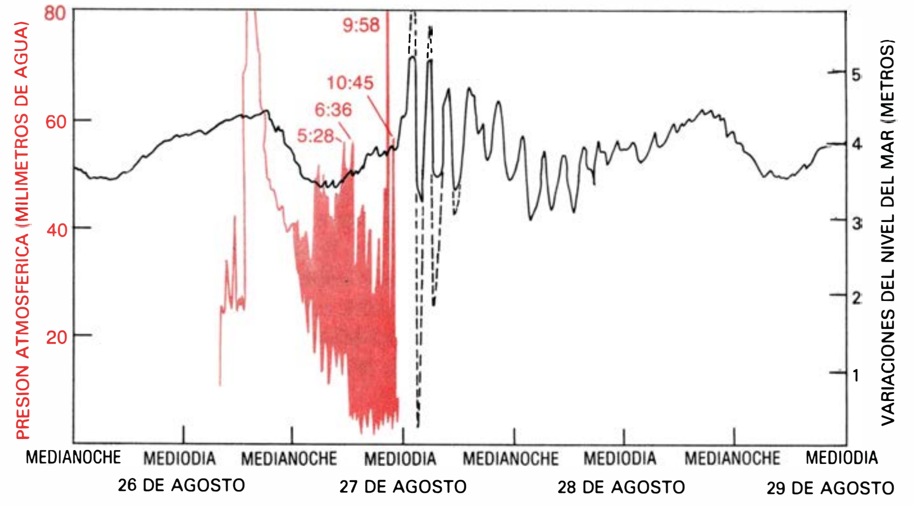


9. MAPA ODOGRAFICO basado en los tiempos de llegada de los tsunamis, registrados por mareógrafos instalados en distintos puntos del estrecho de la Sonda. Hubo que separar primero los tiempos de llegada de los tsunamis de las llegadas de trenes de ondas secundarias producidas al penetrar aquellos en las bahías donde se encontraban la mayoría de los mareógrafos. Se trazó luego su recorrido remontando en el tiempo; se determinó la velocidad en cada instante a partir de la profundidad del mar por el que avanzaban. Las curvas (*trazos de color*) se obtuvieron a partir de las trayectorias y tiempos de recorrido de varios tsunamis (*trazos grises*), lo que permitió establecer el tiempo de llegada de uno cualquiera desde Krakatoa hasta un punto de la costa. Las áreas inundadas aparecen en color. (Los dibujos son de Todd Pink.)

agua. Segundo, existe una probada dificultad a la hora de correlacionar la secuencia eruptiva con las perturbaciones atmosféricas registradas por el manómetro de la fábrica de gas de Yakarta y las olas registradas por los mareógrafos situados a lo largo del estrecho.

Muchas de las olas, especialmente las que se registraron en puntos alejados del volcán, parecen haberse propagado con extraordinaria celeridad, más cercana a la velocidad de las ondas atmosféricas que a las velocidades típicas de las olas marinas. La velocidad de propagación de una ola en el agua es proporcional a la profundidad del medio. En concreto, la velocidad de la ola es igual a la raíz cuadrada de la aceleración de la gravedad multiplicado por la profundidad del agua. Las olas de Krakatoa alcanzaron Honolulú en sólo 11 horas, una velocidad de propagación que implicaría una profundidad media de 17 kilómetros. La profundidad media del fondo del océano es mucho menor: unos cuatro kilómetros. Además, se observaron olas atribuidas a la erupción en partes del océano donde no había justificación racional para esperar su aparición. Llegaron olas hasta el lado opuesto de cadenas de islas, barrera que no podían haber superado.

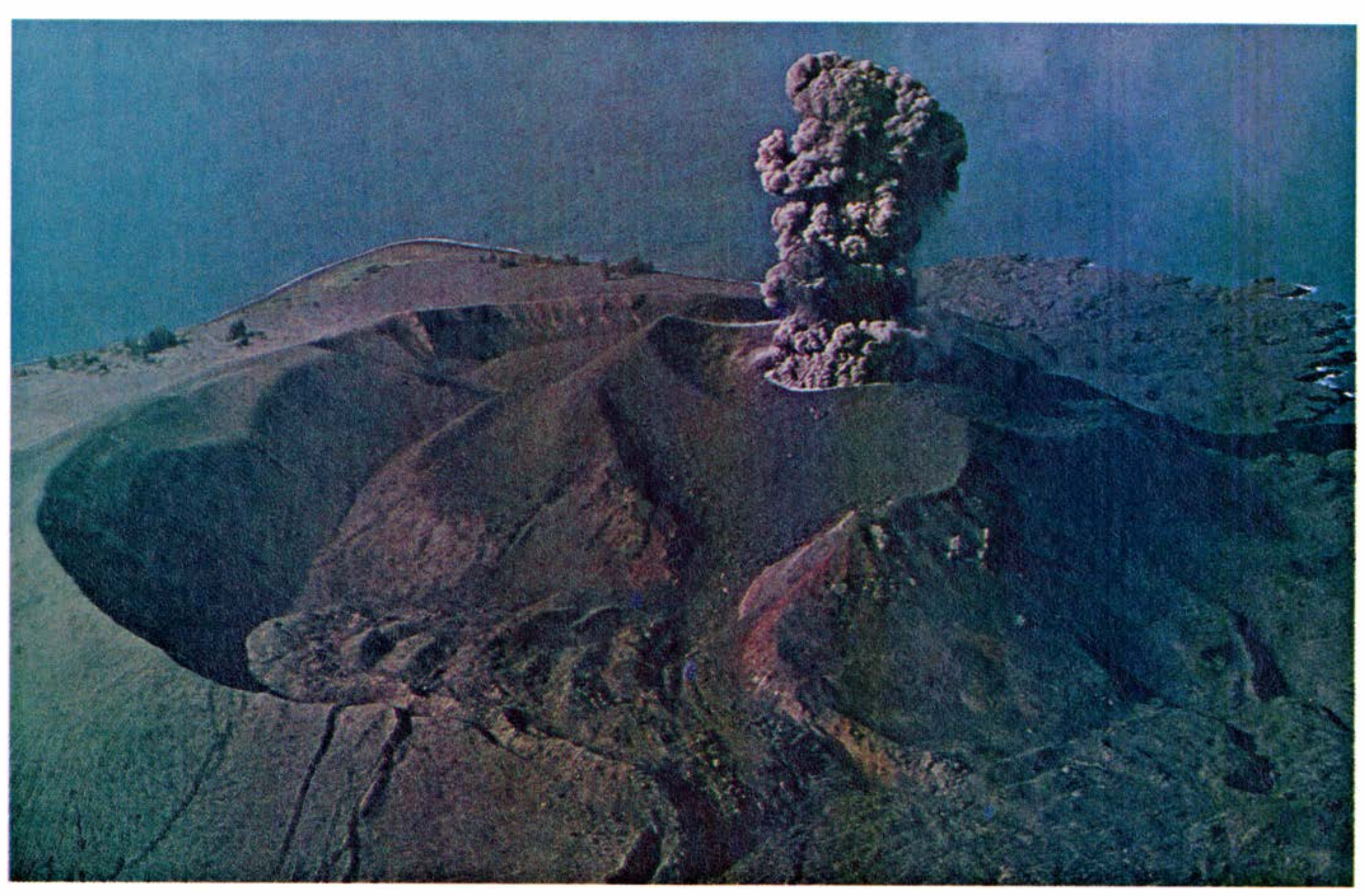
Estas anomalías se apreciaron poco después de la erupción, pero no recibieron explicación hasta muchos años más tarde. En 1955, al disponerse de microbarógrafos capaces de registrar pequeñas variaciones de presión atmosférica, Maurice Ewing, del Observatorio Geológico Lamont-Doherty, y Frank Press, del Instituto de Tecnología de Massachusetts, demostraron que las ondas sísmicas transmitidas por tierra firme podían acoplarse a la atmósfera. En 1967, David G. Harkrider, de la Universidad Brown, y Press demostraron que los largos trenes de pulsaciones de presión producidas en la atmósfera por explosiones nucleares transmitían parte de su energía al océano, levantando olas en regiones distantes de la explosión.



10. LOS MAREOGRAFOS Y UN MANOMETRO DE YAKARTA registraron las principales explosiones y los subsiguientes tsunamis ocurridos el 27 de agosto de 1883. Un manómetro de la fábrica de gas (*color*) registró fortuitamente la llegada de las ondas de presión originadas por las explosiones. El gran pico del registro del 26 de agosto no guarda relación con la explosión; pudo producirse por un aumento de la presión de gas en el depósito controlado por el manómetro. El mareógrafo de Tanjoung Priok (puerto de Yakarta) (*negro*) registró la llegada de las olas levantadas por las mismas explosiones. Algunas de las olas, violentísimas, excedieron el alcance del aparato. Los picos de trazo discontinuo son las estimaciones del desplazamiento del nivel del mar basadas en varias observaciones, tales como el número de peldaños de una escalera que salía del fondo del puerto que quedaron cubiertos por el agua. El mayor de los picos corresponde al tsunami que, generado por la explosión de las 9:58, alcanzó Yakarta a las 12:16. Los picos siguientes incluyen olas secundarias levantadas por el tsunami en el puerto y en el estrecho.

El mecanismo de acoplamiento es complejo; la cantidad de energía transferida depende fundamentalmente de la resonancia, es decir, de las frecuencias de vibración natural de la atmósfera y del océano. No obstante, el principio general que rige la interacción puede expresarse de manera simplificada como sigue: puede considerarse que el océano es sensible a los aumentos y disminuciones de presión atmosférica de tal forma que tiende a mantener el equilibrio hidrostático. Los aumentos de presión atmosférica hacen que descienda la superficie del océano y, las disminuciones de presión, la elevan. Así, el peso de la columna océano-atmósfera sobre cada unidad de superficie se mantiene lo más constante posible. Harkrider y Press reconocieron la analogía entre una explosión nuclear y otra volcánica e indicaron que el mismo tipo de fenómeno explicaría los tsunamis anómalos que siguieron a la erupción de Krakatoa.

Hubo que acometer una paciente investigación para determinar la causa de



11. ANAK KRAKATOA ("Hijo de Krakatoa"), nuevo cono volcánico que se levanta en la caldera creada por al erupción de 1883. Emergió del agua en enero de 1928. Desde entonces, el Anak Krakatoa ha rellenado gran parte de la caldera dejada por la erupción de 1883. (Fotografía de Maurice Krafft.)

los tsunamis que inundaron los litorales del estrecho de la Sonda. Debía discriminarse la llegada de las olas, registrada por los mareógrafos en varios puntos a lo largo del estrecho, de la llegada de olas secundarias levantadas por los tsunamis en las zonas cerradas, como las bahías, precisamente donde estaban instalados la mayoría de los mareógrafos. Había que trazar luego el recorrido de los tsunamis desde su punto de origen, teniendo en cuenta las refracciones, o cambios de velocidad, debidas a la variación de profundidad del agua en el estrecho. En 1981, I. Yokoyama, del Observatorio Volcánico Usu, en Japón, analizó la refracción de las olas de Krakatoa y publicó un mapa odográfico (del tiempo de desplazamiento) de los tsunamis. Basándose en este mapa, llegó a la conclusión de que algunos de los menores se produjeron por la penetración de las proyecciones volcánicas en el mar, pero que el mayor de los tsunamis, que alcanzó Yakarta a las 12:16 horas del 27 de agosto, debió seguir a una erupción submarina.

John H. Latter, del Departamento de Investigación Científica e Industrial de Nueva Zelanda, arribó a conclusiones algo diferentes en un artículo publicado ese mismo año. Empezó por establecer una cronología de los acontecimientos lo más exacta posible partiendo de los registros del manómetro de la fábrica de gas de Yakarta y del mareógrafo del puerto de la capital indonesia. Calculó los tiempos de las explosiones de Krakatoa a partir de los tiempos de llegada de las ondas de presión atmosférica a Yakarta, teniendo en cuenta los ocho minutos que invierten en el trayecto y los cinco minutos de diferencia horaria entre Yakarta y Krakatoa. Entonces, basándose en el cronometraje de las explosiones y en el mapa de tiempos de desplazamiento de tsunamis obtenido por Yokoyama, según el cual el tiempo que tardaba un tsunami desde Krakatoa hasta Yakarta era de dos horas y 25 minutos, demostró que el cronómetro del mareógrafo de Yakarta estaba desajustado en tres minutos y medio respecto al manómetro. Establecida la cronología de los hechos, Latter demostró que los tiempos de llegada a Yakarta de las ondas atmosféricas y de las olas producidas por algunos de los fenómenos de la secuencia eruptiva se correlacionaban.

Pero había casos en que la ola marina no podía correlacionarse con la correspondiente onda atmosférica ni con el acontecimiento explosivo desencadenante. La mayor onda atmosférica, ocasionada por la gran explosión que alcanzó a oírse por todo el océano Indico, arribó a Yakarta a las 10:08 horas. Considerando la velocidad de propagación en la atmósfera y otro factores, debió producirse en Krakatoa por una explosión a las 9:58. El mareógrafo de Yakarta registró una gran ola casi exactamente al mismo tiempo que la onda de presión atmosférica alcanzaba la fábrica de gas. Dado que el tiempo de desplazamiento de las olas desde Krakatoa hasta Yakarta viene a ser de un par de horas y 25 minutos, y que Latter no pudo obtener pruebas de un acontecimiento que ocurriese en Krakatoa a las 7:40, concluimos que la gran ola se produjo por acoplamiento de las ondas atmosféricas con el agua. El verdadero tsunami ocasionado por la explosión de las 9:58, que tenía una amplitud mucho mayor que el falso, no llegó a Yakarta hasta las 12:16 horas.

A esta altura del razonamiento, los detalles cronométricos revisten capital importancia. Dada la velocidad de propagación de la ola, el gran tsunami

que llegó a las 12:16 se habría originado en Krakatoa hacia las 9:45, es decir, antes de que tuviese lugar la mayor explosión. Latter estaba convencido de la precisión de su cronometría, meticulosamente investigada, y estaba seguro de que podía haber una relación causal entre grandes ondas atmosféricas y olas, y por ello llegó a la conclusión de que el tsunami no se originó en Krakatoa, sino en un punto más próximo al puerto de Yakarta, en una distancia correspondiente al tiempo de diferencia. Utilizó entonces el mapa de Yokoyama para demostrar que el punto de origen del tsunami generado por la gran explosión de las 9:58 estaba a unos 10 o 15 kilómetros de distancia de Krakatoa. Tal punto corresponde aproximadamente al margen alejado de la nueva isla de Calmeyer. Otros análisis cronométricos de explosiones y de las olas emparejadas llevaron a Latter a la conclusión de que por lo menos tres de los tsunamis de Krakatoa los debieron provocar las grandes coladas piroclásticas que avanzaban por el mar, similares a la que formó la isla de Calmeyer.

Aunque los tsunamis de Krakatoa son, que se sepa, los mayores de los generados por coladas piroclásticas, ese mismo mecanismo se ha demostrado responsable de tsunamis que siguieron a la erupción de otros volcanes. En 1980, Juergen Kienle y Samuel E. Swanson, del Servicio Geológico de los Estados Unidos, comunicaron que el tsunami de nueve metros que barrió completamente English Bay, Alaska, después de la erupción del volcán Augustine en 1976, lo desencadenó una colada piroclástica que penetró en la bahía.

Un nuevo volcán, Anak Krakatoa ("Hijo de Krakatoa") se ha levantado aproximadamente en el lugar que ocupaba el antiguo cono de Danán antes de la erupción de 1883. Se trata de un volcán antaño submarino, cuyo cono emergió a finales de enero de 1928. Desde entonces, sucesivas erupciones han ido llenando gradualmente con lava y coladas piroclásticas la porción septentrional de la caldera de 1883. El foco de actividad volcánica se está desplazando hacia el sur; probablemente siga la misma fisura que determinó la alineación de los conos volcánicos en la isla primitiva. Quizá termine por formarse una gran isla alineada como la vieja, de noroeste a suroeste. Una nueva Krakatoa renace, como el ave fénix, de las cenizas de sus restos.

Juegos de ordenador

Reseña de actualidad sobre el noble arte de transformar literatura en parloteo

Brian Hayes

Suministrándoles información lo suficientemente embrollada, se puede conseguir que casi todos los programas de ordenador produzcan resultados absurdos. No es otro el sentido de la fatigada moraleja "Da basura y te la devolverán más pura". Tan sólidamente fundamentado está hoy el dicho, que nadie prestaría mucha atención a nuevas comprobaciones. Empero, con algún esfuerzo y un poco de reflexión, podemos conseguir programas que admitan como entrada obras maestras de la literatura y generen como salida las bobadas más palmarias. Allá va el tercer acto de *Macbeth*; salir, sale un bodrio narrado por un idiota, repleto de sonido y furia, pero que nada significa. Ahora bien, también *esto* es procesar datos. (La transformación inversa, por desgracia, es más difícil.)

El proceso de transformación de un texto literario en un galimatías se efectúa en dos etapas. Primero, el programa "lee" el texto, extrayendo y registrando ciertas propiedades estadísticas. Estos datos estadísticos definen la probabilidad de que una determinada letra cualquiera del alfabeto vaya a continuación de otra letra, o de otra secuencia de letras, del texto fuente. En la segunda fase se procede a generar un nuevo texto, eligiendo las letras al azar en concordancia con las probabilidades registradas. De aquí resulta un río de caracteres que reproducen las propiedades estadísticas del texto original, pero cuyo significado, de tener alguno, es meramente accidental.

Es difícil imaginar un método de imitación más burdo. En ningún punto del programa hay tan siquiera una representación de la noción de palabra, y no digamos ya indicio alguno de lo que las palabras pudieran significar. No hay representación de ninguna estructura lingüística más elaborada que la mera sucesión de letras. El texto así creado es un *pastiche* del tipo más desmañado y torpón, que conserva tan sólo las más superficiales cualidades del texto original. Lo notable es que el producto de este simple ejercicio nos resulte a veces de una misteriosa familiaridad. Es un resultado absurdo, pero no un absurdo indiferenciado. Es más bien un absurdo de timbres shakespearianos, chaucerianos o jamesianos. De hecho, es como si, al eliminar todo el contenido semántico, las peculiaridades estilísticas se hicieran tanto más conspicuas. Lo que nos hace preguntarnos cuán próximos a la superficie son los rasgos que definen el estilo de un autor.

El proceso de generación aleatoria de prosa ha sido investigado con detalle por William Ralph Bennett, Jr., de la Universidad de Yale. Bennett ha hecho de la estadística del lenguaje uno de los temas principales de un curso sobre las aplicaciones de los ordenadores; el mismo tema ocupa lugar prominente en su libro de introducción a la programación, *Scientific and Engineering Problem-Solving with the Computer*. (El libro es mucho más ameno de lo que su título podría sugerir. Entre los problemas en él examinados se encuentran los de la aerodinámica del partido de rugby entre los equipos de Princeton y Dartmouth de 1950, que se jugó durante un huracán; la difusión de la sífilis en una población de marineros y prostitutas y el análisis espectral del oboe, el *krummhorn* y la "manguera del jardín bien temperada".)

Bennett hace notar que la más antigua referencia a la generación aleatoria de lenguaje de que se tenga conocimiento se encuentra en *Maxims and Discourses*, de John Tillotson, arzobispo de Canterbury entre 1690 y 1700, en defensa de la creación divina del mundo. Tillotson escribió: "¿Cuántas veces tendrá un hombre, habiendo mezclado en una bolsa una colección de letras, que arrojarlas al suelo para que formen al caer un poema exacto, o bien, hasta formar un buen discurso en prosa? ¿Y no será más fácilmente creado por azar un librito que escrito el inmenso volumen que es el mundo?"

Para casi todas las consideraciones modernas acerca del lenguaje creado por azar, el punto de partida es un pasaje de Sir Arthur Eddington, que data de 1927. "Si pusiéramos a un ejército de monos a aporrear máquinas de escribir, *podrían* llegar a escribir todos los libros del Museo Británico". También Eddington se proponía cargar las tintas en lo muy inverosímil de semejante resultado; lo citaba como ejemplo de acontecimiento teórico que jamás se daría en la práctica. Sea como fuere, desde los tiempos de Eddington la posibilidad de descubrir el genio creador entre los aleatorios aporreos de los monos ha cobrado literariamente vida propia. Bennett menciona al respecto obras de Russell Maloney y Kurt Vonnegut, Jr., así como un espectáculo de club nocturno creado por Bob Newhart.

El proceso que Eddington imaginaba puede ahora simularse mediante un programa que yo llamaré generador de textos de orden cero. Ante todo, es preciso convenir un alfabeto, o juego de caracteres, que especifique qué teclas deberán instalarse en las máquinas de escribir que los monos habrían de aporrear. En ciertas simulaciones de orden superior resulta importante que el número de caracteres sea el menor posible y, por razones de consistencia, lo mejor parece ser adoptar el mismo sistema de caracteres que para el programa de orden cero. He seguido, así pues, las recomendaciones de Bennett, y elegido un sistema de 28 símbolos: las 26 letras mayúsculas, el espacio de separación entre palabras (que el ordenador trata como un carácter más) y el apóstrofe (que es en inglés escrito más corriente que las tres o cuatro letras menos utilizadas).

Un mono hipotético ideal, carente de sesgos y preferencias, tendría en cada momento igual probabilidad de pulsar cada una de las teclas. Es fácil simular tal conducta. Podemos asignar a cada símbolo del sistema un número, desde 0 a 27. Para cada carácter a generar se elige al azar un entero de esa gama, y se imprime el correspondiente carácter. Vemos en la figura 1 un breve espécimen de texto generado por este proce-

```
'PWGMMLTHIDVGRHPEDFCXFEKFNOPYPQSXZRUXG'YS'AEEU PEDEGLQYFUWPO'IKI
QTONIXJKZEUKDXWKKJREHYHPKWUJHLEJNBPLQ AIEOQXUBJYYVIFFDPQGIGZNTI
RQXPDJ NQESPQMCRSNGMKQEZICZV'GSWALK ZZEYIBBOTDCRSMK'VI MRCZXUBI
SNEQ'VQQHFQUCBJXZRVVNIBHFJEFTCFJPWFOIYHOMPNFSFWKNCMVLOJJBX
QV KIZTLNRWGGTZFPZPQQCGVJCPAYRDQJRMYSWCGABRXLERCYYRHQCHTOQ'UT
FMRITFTIZUIWTSTXWQGOCAFXJOZYKSTV'BYOBEUFIRQWQ VOUVQJPRKJWBKPLQZCB
```

1. Texto aleatorio de orden cero, con alfabeto de 28 símbolos

dimiento. No tiene parecido ninguno con el inglés escrito, ni con ningún otro lenguaje humano. Sus "palabras" tienden a ser excesivamente largas (veintisiete caracteres, por término medio), y recargadas de consonantes. El motivo, como es obvio, es que en los idiomas naturales las frecuencias de las letras distan mucho de ser uniformes. El espacio inter-palabras cuenta en ellos alrededor de la quinta parte de los caracteres, mientras Q, X, W y Z apenas si alcanzan conjuntamente a sumar el uno por ciento. En una simulación de orden cero, todos los caracteres tienen la misma frecuencia, a saber, 1/28.

El número cómico de Bob Newhart se refiere al sino de los inspectores que han de leer la producción literaria de los monos. Tras muchas horas de escudriñar ininteligibles parloteos, atinan con la frase "To be or not to be, that is the gesorenplatz...". La verdad es que incluso llegar a tanto es sumamente improbable; no es de esperar que las nueve primeras palabras del soliloquio de Hamlet se presenten más de una vez cada 2×10^{46} caracteres. En una pasada de 50.000, yo tan sólo pude encontrar un ejemplo de TO y otro de NOT, y separados además por muchas líneas. (Desde luego no fui yo quien leyó los 50.000 caracteres, sino que efectué la búsqueda con un programa encargado de detectar pautas.)

PRIMER ORDEN

HUD T ALONIT NTA SN TVIOET ELERFOAD PE TRLTWTL N CABEG TYLUEMU TIGT
BH OFDRRIC O STU HOOOTO YATNDL UYA HWAE SS NLSDB OTRORT DEERARFT
D LBFF HHARE MW OSPE OFOIT SEOUN GTUMG H N GHKOY T EAOS A SD E TNNE
PEHAGIADIHNATO AATSAGI ED INNE ABRA TAAM GT E TWNO HEWIIGUTNCM GA SFHHY
HREBH RARE OOSY LFE OC EGGTA WIFRTYE EUS DA ETO WF EIT ERNETEBTSTTELO
NTAAN O YEETWNSONRNHN TYHVN NLUESETTHLGEAKPNNMTIA TSM REEANTVONC POE
RUTP EOIT L IEETGTWHSW H KHHER W OLIOEWOEPT D AEYBSTNHGDNPT C TNLINHH
KHHE E RTVIOB EI K EOAFPUTSTTAS NA LAN SRDF D NMTHESKO UGEEDICRAWDT OBD
TUIML WSORGNETE

SEGUNDO ORDEN

BEGASPOINT IGHIANS JO HYOUD WOUMINN BONUTHENIG SPPRING SBER W IDESE WHE D
OOFOMOUT O CHEDA AFOOIAUDO IS WNY UT DRSASER LD OT POINE ETHAT FOEVEL BE
ORRI IVER BY HE T AS I HET W BE T WAU GIM UTHENTOTETHAVE THIKEWOITOCOUTORE
TATHASTHEE AT D Y WAN TOND SE TEDING US AKIN WING W TE T BO TOTSTHINGATONO
EN T LLY WID OUCOUSIND HEF THIMES AG T BENG LORYE ALLATHOMOFTHER TOUDIMS YS
S ORYRY THERNG S HE M G M ANG S CITOOFO HEN G BEST ONDLOL ANE DO HE
ICISEKERIT ME NKITHADIMUPL WHES HT BATHE T LOR WITULOWAYE WATHEG M
LEROMAUN OUGS POUPO O HASING LIN ON ASHAN AWFAS HET ND MEDE

TERCER ORDEN

MAY THOT TO THER YOURS CHIM JOSE EY EILLY JUSED AND HID YEL THE MARK WASK
TROOFTEN HEREY LING SH THAVERED HER INCED I MEA BUT DAY WOM THE EAKIN WIPS
AS SUGH THE WAY LIARADE TH MY HE ALMASEETIR ANICIOUT JOSIDNTO GRATEVE NO
VER BIGH WER ACCOW WAS I GEORE HENDSO EGGET PUT TO SQUAD TRADE OFF GIN
GO ME HER SPING HE CONE WELL FEWHEY THEYES AND AND QUICE YOULDNT HER
ORL SO MAKING RINGS SOMET DREAVE HISETTO COMAD THAT ME WE MIG TOLD THE
THERFUMBECK OT OFF FEELP HE WAST ITS LETHOTTEN ITHEE ROWN YOURS FEL FOR
SOME IF WIS HE STAKED UPPOIS SHENS NO TILL HIM I WAY SO WHATEALWAS WER TWE
NER DING O THIS IT IN ANIGH ACK REAN THAT DO GETHE BITER

2. Textos aleatorios de primero, segundo y tercer orden, basados en el soliloquio de "Molly Bloom"

Un primer paso encaminado a perfeccionar la capacidad literaria del simio escritor consiste en ajustar la selección de cada letra, de modo que refleje la frecuencia real con que tal letra aparece en el lenguaje escrito. El plan consiste, en efecto, en construir un teclado con, pongamos por caso, 2500 teclas espaciadoras, 850 letras E, 700 teclas T, etcétera. Las frecuencias de las letras pudieran ser promedios calculados a partir de un largo texto en prosa, en el idioma de que se trate, pero resulta más adecuado e interesante basarlas en promedios extraídos de un texto fuente determinado. Un programa que seleccione caracteres de acuerdo con una tal distribución de frecuencias será llamado "generador de textos de primer orden".

Los valores de frecuencias de las letras pueden quedar representados en una lista (variable dimensional, o "registro") unidimensional de 28 elementos. El registro es un bloque de alojamientos para datos, que se reserva en la memoria del ordenador, y está organizado de modo que cada elemento queda especificado por un índice, comprendido entre 0 y 27. Para ir creando esta lista podríamos contar cuántas veces aparece cada letra en el texto fuente, y cargar los valores a mano. Es preferible, sin embargo, que sea el propio programa quien se encargue del recuento, aun cuando ello implique preparar el texto fuente para que le sea "legible". El programa de recuento comienza haciendo iguales a cero todos los elementos del registro. El texto es entonces inspeccionado, a razón de un carácter por vez; la lectura de un carácter conlleva el incremento en una unidad del contenido del elemento del registro a él correspondiente.

El texto aleatorio de primer orden se genera haciendo que la probabilidad de elección de un determinado carácter sea proporcional al valor del elemento del registro asociado al mismo. He aquí un posible método: se genera un número aleatorio en el intervalo comprendido entre cero y una cota superior igual a la suma de los elementos del registro, o lo que es igual, al número total de caracteres del texto fuente. El primer elemento del registro, que pudiera reflejar el número de apariciones de la letra A, se resta del número aleatorio. Si el resultado es menor o igual que 0, se imprime una A; de lo contrario se sustrae el elemento siguiente (que representa el número de B) del valor restante tras la primera comparación. Se continúa de igual modo hasta la primera sustracción que dé negativa o nula, quedando seleccionado entonces el correspondiente carácter. Obsérvese que el procedimiento no puede dejar de elegir alguno de los caracteres, pues el número aleatorio no puede exceder de la suma de los elementos del registro.

En la figura 2 vemos una muestra de texto aleatorio de primer orden. Está basado en una lista de frecuencias compilada de un pasaje del último capítulo del *Ulysses* de James Joyce, el conocido por "Ithaca", o soliloquio de Molly Bloom. Para elegirlo tenía yo un motivo: la falta de signos ortográficos característica del texto aleatorio carece de importancia, al no tenerlos tampoco el texto fuente.

Aunque la información que sobre frecuencias de las letras queda incorporada a los textos aleatorios de primer orden comporta cierta mejora, malamente podríamos decir que el texto es legible. Si bien la longitud promedio de las palabras (4,7 letras) se aproxima al valor esperado (4,5 letras), la varianza, o la desviación del promedio, es, con mucho, demasiado grande. Al parecer, las palabras del inglés ordinario no sólo son breves, sino que sus longitudes oscilan en una estrecha franja. En el texto aleatorio la distribución es mucho más amplia. Aparte de la cuestión de la longitud de las palabras, tenemos el problema de su contenido. Aunque las letras aparezcan con las frecuencias correctas, la sucesión en que se presentan es manifiestamente aleatoria, y la mayoría de las "palabras" resultantes, ni son inglesas, ni lo podrían ser. Una

serie de letras como WSTLNTTWNO, O HIUOIMYTG no sólo carece de significación: es imposible. En una pasada de 2000 caracteres, la más larga de las palabras reconocibles fue, muy apropiadamente, RARE.

El refinamiento siguiente representa el paso crucial, pues podría continuarse, al menos en teoría, hasta órdenes arbitrariamente altos. La raíz de la idea es que la probabilidad de que en el idioma escrito aparezca una letra en un lugar dado depende fuertemente de las letras precedentes. Por ejemplo, tras una V es seguro que vendrá una vocal; tras una Q ha de venir forzosamente una U. El procedimiento consiste, pues, en formar una tabla de frecuencias individual para cada uno de los símbolos del conjunto de caracteres. Las frecuencias son entonces registradas en una tabla bidimensional de 28 filas y 28 columnas, con un total de 784 elementos. Vemos un ejemplo de una tal tabla en la figura 4. (La tabla está "normalizada" por filas, lo que significa que las comparaciones sólo son válidas dentro de cada fila.)

Cuando se genera un texto a partir de esta tabla bidimensional, el último carácter elegido determina qué fila de la tabla debe examinarse para seleccionar el carácter siguiente. Por ejemplo, si la letra precedente fue una B tan sólo se toman en consideración los elementos de la segunda hilera. El mayor de los elementos de esta fila es E, que es por tanto la letra de mayor probabilidad. También A, I, L, O, R, S y U tienen cierta probabilidad de resultar elegidos. Las combinaciones imposibles, como BF y BQ tienen frecuencia cero y, por tanto, jamás aparecerán en la salida del programa.

El texto aleatorio de segundo orden comienza a mostrar los primeros síntomas de verdadera estructura lingüística. La distribución de las longitudes de las palabras es tan sólo algo mayor de lo que debiera. No es excepcional encontrar auténticas palabras, y hay además muchas "vice-palabras" (tales como SPPRING O THIMES); una gran proporción de las series de letras son, cuando menos, pronunciables. Comienzan a aparecer frecuentemente los dígrafos ingleses más comunes, como TH; la alternancia de vocales y consonantes va siendo detectable ya.

El paso siguiente debiera ahora resultar obvio. Un algoritmo de tercer orden elige cada letra del texto aleatorio de acuerdo con las probabilidades que determinan las dos letras precedentes. Lo cual requiere una formación tridimensional de 28 planos, cada plano formado a su vez por 28 filas de 28 columnas. Supongamos, por ejemplo, que haya sido generado el grupo TH. El programa debe entonces dirigirse al vigésimo plano (correspondiente a la T) y a la octava fila de ese plano (correspondiente a una H). En ese fila la elección más probable es E, aunque A, I, O y el símbolo de espacio blanco tienen también una posibilidad. De seleccionarse realmente la E, en la iteración siguiente se procederá a elegir la quinta línea del octavo plano, que es la porción de tabla especificada por la secuencia HE. Aquí, el más distinguido candidato es el espacio de separación entre palabras, seguido por R.

AD CON LUM VIN INUS EDIRA INUNUBICIRCUM OMPRO VERIAE TE IUNTINTEMENEIS
MENSAE ALTORUM PRONS FATQUE ANUM ROPET PARED LA TUSAQUE CEA ERDITEREM IN
GLOCEREC IOVELLUM ET VEC IRA AE DOMNIENTERSUO QUE DA VIT INC PARBEM ETUS
TU MEDE DERIQUORUMIMO PEREPORIDEN HICESSE COSTRATQUIN FATU DORAEQUI POS
PRIENS NOCTA CIENT HUCCEDITAM PET AUDIISEDENDITA QUE GERBILIBATIA VOLAEQUE
ORECURICIT FES ADSUE ARCUMQUE LULIGITO PIMOES PERUM NOSUS HERENS EA
CREPERESEM ETURIBUS AVIS POS AT IS NOMINE FATULCHENTURASPARIS AUDEDET PARES
EXAMENDENT DUM REMPET HA REC ALEVIREM ORBO PIERIS ATAE PARE OCERE RAS

QUALTA 'L VOL POETA FU' OFFERA MAL ME ALE E 'L QUELE ME' E PESTI FOCONT E 'L M'AN
STI LA L' ILI PIOI PAURA MOSE ANGO SPER FINCIO D'EL CHI SE CHE CHE DE' PARDI
MAGION DI QUA SENTA PROMA SAR OMI CHE LORSO FARLARE IO CON DO SE QUALTO
CHE VOL RICH'ER LA LI AURO E BRA RE SI MI PAREMON MORITA TO STOANTRO FERAI TU
GIA FIGNO E FURA PIA BUSCURA QUAND'UN DEL GUARDI MIN SA PAS DELVENSUOLSI PER
MUSCER PIE BRUI TA DORNO TITTRA CHE PO E PER QUE LI RINONNIMPIAL MIN CH'I'
BARVEN TA FUI PEREZZA MOST' IO LA FIGNE LA VOL ME NO L'E CHE 'L VI TESTI CHE
LUNGOMMIR SI CHE FACE LE MARDA PRESAL VOGLICESA

PONT JOURE DIGNIENC DESTION MIS TROID PUYAIT LAILLE DOUS FEMPRIS ETIN
COMBRUIT MAIT LE SERRES AVAI AULE VOIR ILLA PARD OUR SOUSES LES NIRAPPENT LA
LA S'ATTAIS COMBER DANT IT EXISA VOIR SENT REVAIT AFFRUT RESILLESTRAIS TES FLE
LA FRESSE LES A POURMIT LE ELLES PLOIN DAN TE FOLUS BAIER LA COUSSEMBREVRE
DE FOISSOUR SOUVREPIACCULE LE SACTUDE DE POU TOUT HEVEMMAIT M'ELQU'ILES
SAIT CHILLES SANTAIT JOU CON NOSED DE RE COMMEME AVAIL ELLE JE TER LEON DET
IL CED VENT J'ARLAMIL SOUT BLA PHYSIS LUS LE SE US VEC DES PEUSES PAU HAS BEAU
TE EMANT ELLE PLANQ HEUR COIRACOUVRE BIENE ET LUI

3. Latín (Virgilio), italiano (Dante) y francés (Flaubert) de tercer orden

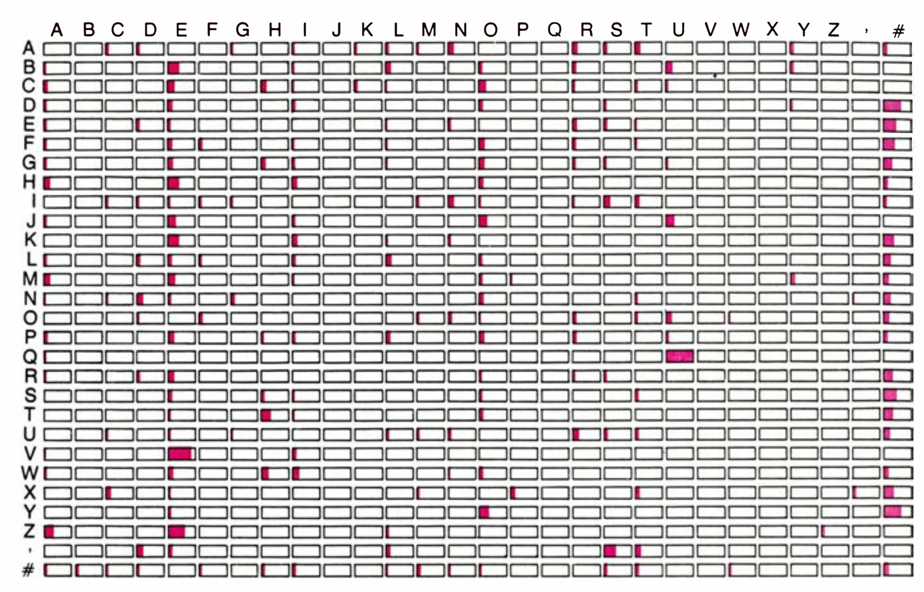
En el texto de tercer orden no pueden aparecer secuencias de tres caracteres a menos que se encuentren presentes en algún lugar del texto fuente. Dado que en el recuento se atiende a los espacios en blanco, ello basta para garantizar que todas las palabras de una letra sean auténticas palabras. El rendimiento práctico es, sin embargo, muy superior al garantizado: virtualmente todas las secuencias de dos letras son palabras, y otro tanto vale para casi todas las secuencias de tres letras. No es raro que se presenten secuencias de varias palabras: PUT TO SQUAD TRADE OFF GIN GO ME HER. Incluso ciertas secuencias largas, aunque no llegan a formar palabras, tienen cierta plausibilidad fonética. Después de todo, sólo por accidente carece de significación en inglés ANYHORDANG HOUP TREAFTEN.

Al leer muestras de texto aleatorio de tercer orden no puedo evitar pensar en las imitaciones que de otros idiomas se hacen en el teatro, o en la "glosolalia", el don de lenguas de que se nos habla en la liturgia de Pentecontés. Podría conjeturarse que el parecido es síntoma de algo; tal vez la gente que ha adquirido tales artes efectúa inconscientemente un análisis estadístico, un poco a la manera que lo hace un ordenador. Se me ocurre otra explicación más verosímil: la imitación de otros idiomas y la glosolalia parecen implicar un ensamblaje aleatorio de los átomos fundamentales del lenguaje oral, los fonemas. Pudiera ser que el tamaño idóneo de las representaciones escritas de los fonemas rondase en torno a tres letras.

En el texto de tercer orden, las peculiaridades estilísticas del texto fuente comienzan a mostrar efectos perceptibles. Cuando el contraste de estilos es grande, los correspondientes textos aleatorios son también claramente distinguibles, si bien no es fácil decir en qué reside exactamente la diferencia. Me siento inclinado a describirla como una cuestión de textura, pero no estoy nada seguro de lo que sea textura de una prosa. ¿Será lo que pueda subsistir al eliminar del texto todo significado?

No es posible percibir en la prosa aleatoria de tercer orden los idiotismos del autor del texto fuente, pero no hay dificultad en reconocer su idioma. Las pautas de vocales y consonantes, y las terminaciones de las palabras, características, son inconfundibles. La figura 3 nos muestra ejemplos breves de latín (Virgilio), italiano (Dante) y francés (Flaubert). Quien sólo conociera de "aspecto" alguno de estos idiomas pudiera tener dificultad en distinguir el sucedáneo del auténtico.

Antes de examinar lo que pueda haber más allá de las aproximaciones de tercer orden me gustaría mencionar algunas otras aplicaciones de las tablas de frecuencia de letras. Bennett, analizando la entropía del lenguaje, señala que las tablas nos permiten calcular la cantidad de información aportada por cada carácter del texto. El contenido de información mide, fundamentalmente, la dificultad de predecir el próximo carácter de un mensaje. Alcanza un máximo en la simulación de orden nulo, donde todos los posibles caracteres tienen iguales probabilidades; dicho de otra forma, el contenido de información es máximo cuando el texto es totalmente ininteligible. La noción de predictibilidad de los caracteres conecta con el análisis de corrección de errores en telecomunicaciones y con el diseño de algoritmos para resolver textos cifrados y criptogramas.



4. Tabla de frecuencias de segundo orden correspondiente al Acto III de **Hamlet**

Otra región digna de ser explorada es la alteración o manipulación de la tabla de frecuencias. ¿En qué medida cambia el texto aleatorio si, por ejemplo, cada elemento de la tabla se eleva al cuadrado? Un ejemplo de "cuadratura" de Molly Bloom se muestra en la figura 5. Dado que el procedimiento exagera las diferencias entre los elementos de la tabla, el efecto es el de "contrastar" más la distribución de frecuencias: las palabras corrientes se hacen más corrientes todavía. Hay muchas otras transformaciones posibles. Al sumar una misma constante a todos los elementos de la tabla se producen resultados desastrosos, aun cuando la constante sea pequeña, pues todas las combinaciones imposibles de letras, que tanto trabajo nos había costado eliminar, resultan nuevamente posibles.

Una idea curiosa es la de multiplicar por −1 la tabla entera, a fin de generar textos de, pongamos por caso, Alexander anti-Pope. Por cada combinación dada de letras, si una letra cualquiera es la más probable siguiente inmediata en Pope, pasará a ser la menos probable en anti-Pope. Sin duda, la habilidad literaria quedaría mejor servida si el producto recordase las obras de Colley Cibber. Pero la realidad es que se obtiene un revoltijo casi totalmente desprovisto de regularidades.

El resultado es algo menos descorazonador (aunque diste mucho todavía de ser deslumbrante) al sumar o multiplicar dos tablas. Podemos así crear obras colectivas por colaboradores de improbable avenencia: Jane Austen más Mark Twain, o bien Keats multiplicado por (Byron más Shelley). Lo que más me gustaría es ver qué resulta de Byron menos Shelley, esto es, la esencia destilada de sus diferencias. Desdichadamente, no he podido hacerlo funcionar. Casi toda la información de una tabla de frecuencias de tercer orden representa la estructura lingüística común a los autores de mismo idioma. Al restar este elemento común apenas si queda otra cosa que ruido.

Hay una razón más fundamental, que explica el fracaso de la diferencia de dos tablas. En las tablas de tercer orden, alrededor del 90 por ciento de los elementos son ceros, que corresponden a la gran mayoría de combinaciones de letras que jamás se observan en inglés, como RJT o UUU. Ordinariamente, el programa no puede jamás "aterrizar" sobre uno de estos elementos, pero en cuanto la tabla es modificada por sustracción resulta casi inevitable topar con una fila cuyos elementos sean todos ceros. No hay forma airosa de salir de esta vía muerta.

No hay dificultad en concebir un programa capaz de crear una tabla de frecuencias y generar texto aleatorio; lo que sí presenta dificultad es encontrar espacio de almacenamiento suficiente para la tabla tridimensional del modelo de tercer orden. Fue la necesidad de hacer mínimo el tamaño de la tabla lo que nos hizo limitarnos a un juego de 28 caracteres. Incluso con esta limitación, la tabla tiene casi 28.000 elementos, y cada elemento puede exigir dos bytes, que son las unidades básicas de almacenamiento. Para alojar la tabla, más los programas necesarios, con la capacidad de memoria de un ordenador pequeño puede hacer falta algo más que calzador.

En el siguiente orden de aproximación cada carácter se selecciona en razón de las probabilidades determinadas por los tres que le anteceden. Se requiere una tabla tetradimensional, con un total de más de 600.000 elementos. En 1977, en un artículo publicado en *American Scientist*, Bennett dio especímenes de textos de cuarto orden, generados tras construir tan enorme tabla. También escribió, esta vez en un libro, que la simulación de cuarto orden "ronda los límites prácticos, con los mayores ordenadores disponibles en el momento actual". Con los pequeños ordenadores a los que la gente pueda tener acceso fácil, el cuarto orden parece definitivamente fuera de alcance.

Mas los "límites prácticos" han sido creados para ser traspasados, y cuando

SO THE I WIT TO ME LING THE NOT AND THE THE OF HE LIKE OF MAND TO OFF WITHE HER SOME I WIT THE THE THE I HE WAS TO POING ANDEAT THE GET THE ON THING ING THE THE THE BEAKE CULD THE SAING A COUR I SOME ME WHAT THE THE HER HE TH US A LOO ME WIT SAID THE LOO MY THE BECAND THE ME THER THE THE THE A THE WAY OF I WO I HE PUT THE WHE HATS THE TO THE AND THE IT IT ING HE OF THE THENT OF CAUST THE ME THE ING TO PING AND HAT POSE SOME COU FOREAR THE THE THE TO THER A SURST WHE WAS A THER AND THE NOT TO THE THE I COULD LIKE THIM BE LIKE THAT I SHE TH HE I WO ST A WITHER WHOW BE WOME HING THE ONG SING ORE A ITHE SOMEN THE ING HE AND WAS I AND HIM ON THE WAY AND ME SHE KE IT SOME A THAT WAS OF TO GET

5. Una tabla modificada origina prosa por "Molly Bloom al cuadrado"

el problema se examina desde otro ángulo, la perspectiva no es tan desalentadora. Como ya hicimos notar, casi todos los elementos de la tabla de tercer orden son ceros; es de esperar que la orden cuarto tenga todavía mayor número de elementos vacíos. Concebí entonces un plan: en lugar de almacenar las frecuencias en una inmensa tabla tetradimensional casi desierta, yo crearía muchas listas unidimensionales pequeñas. Cada una de éstas sería equivalente a una de las hileras de la gran tabla, pero su longitud sería la justa para albergar los coeficientes no nulos. Las líneas cuyos elementos fuesen todos nulos se eliminarían sin más.

El plan es teóricamente factible, me parece, pero engorroso. Asignar espacio de alojamiento a 10.000 o más registros cuyo tamaño puede ir desde 1 hasta 28 elementos parece tarea de feo aspecto. Como se verá, encontré un procedimiento mejor –o al menos, más sencillo– que permite generar texto aleatorio de orden arbitrariamente alto, con un juego de caracteres que abarca el alfabeto entero, más todos los signos ortográficos que el ordenador sea capaz de imprimir o presentar en pantalla. Como era de temer, se paga por ello una penalización: el método es unas 10 veces más lento.

Me vi llevado a tomar en consideración otras alternativas al procedimiento de Bennett fantaseando acerca de los casos extremos del proceso de construcción de registros. Imaginemos un texto fuente, escrito con un alfabeto de 28 símbolos, que esté formado por 10.001 caracteres. La máxima de las posibles tablas de frecuencias que describen su estructura es entonces de orden 10.000. Tiene 10.000 dimensiones y $28^{10.000}$ elementos, número absurdamente grande, para cuyo tamaño son imposibles las metáforas. Es un número, sencillamente, inimaginable. Y lo que es más, de todos esos innúmeros elementos del registro, tan sólo uno es distinto de cero: el que ocupe dentro de la lista el lugar especificado por los 10.000 primeros caracteres del texto, y cuyo valor determina el último carácter. Aun cuando pudiera crearse un registro semejante (y el universo no es lo bastante grande para darle cabida), la idea de tomarse tantas molestias para determinar únicamente a un elemento parece –y es– descabellada.

Con formaciones y listas de orden pequeño la impresión de desproporción es menos extrema, pero sigue siendo perceptible. El hecho es que toda la información que pueda quedar incorporada a cualquier tabla de frecuencias, por grande que sea, se encuentra presente en el texto original, y es en éste donde adopta su forma más compacta. (El razonamiento que sustenta tal afirmación es singularmente difícil de expresar; tiende casi a la tautología. Lo que registra la tabla de frecuencias es la frecuencia de las secuencias de caracteres en el texto, pero esas secuencias, y exactamente esas secuencias, se encuentran presentes en el texto fuente, en exactamente la frecuencia registrada.)

El método para generar texto aleatorio sugerido por esta observación funciona como sigue. Se crea una única tabla de frecuencias; es una formación unidimensional pequeña, compuesta solamente por tantos elementos cuantos símbolos integren el sistema alfabético y de signos de puntuación utilizado. Yo opté por tomar 90 caracteres. El ordenador lee entonces el texto fuente completo y lo aloja en su memoria, almacenado, en el caso más sencillo, como una "serie, o cadena, de caracteres" ininterrumpida (lo que suele conocerse por *string*). A continuación se selecciona una secuencia de caracteres con la que comenzar el texto aleatorio, que yo llamaré secuencia patrón.

La tarea de ir cargando los elementos de la tabla de frecuencias se efectúa revisando por completo el texto fuente, buscando hallar todos los posibles casos donde se presente la secuencia patrón. Por ejemplo, si la secuencia patrón es "gain", en la revisión podrían ser detec-

CUARTO ORDEN

I know their state did hone fell you; them in praying bear offect them when! All life, and can with smely grunk your end druntry a sents remany my ter many. Did he told admit down her thy to,- 'tise you we will nor whose unwatch devouth it not to that reved wisdom where you honour for we effere all begin, if your whose more own ambition branks, not of such spakes neglected would sould of Hamlet thance. To abountry word. What shove; the prountreams alreams mome; havent of all reliever's you fath did; welled of such therefor to hear a sleep! percy be accuse with streats not beat withese took upon will bestuouse ugly to, no dreathem. 'Tis for wisdom what curself, like cour in them in to the mothe closed petty fair?

QUINTO ORDEN

I, his soul, that are. To a nunnery. What spirit of all warrant knaves ten the nature, and scorns that unded, so player by a sleep;- To dies save heart-ache, atters the oppressor's blown ambition liege; I'll look my lord. O heart; and I'll give that he spokes thy origin and love. Her fault is night his fit, and quickly justice, and man's chaste as you now rights. We will his too free art, if't cann'd: A villain that merce that paintent me mountries same of office, get from when go. Oh, 'tis somethings and drift of him in. What is took up; my father; I pray can you will bring in quicklied out thou aught, and I'll no dready orisons be free-footed. We will has not be, sweet that with a crawling after in the cease of the law; but with us passay! Bow, stubborne me my mother aloof, what reply.

SEXTO ORDEN

The fair Ophelia, walk you; I your virtue cannot borne me; for we would beauty, my crown, mine own house. Farewell. Oh, my lord. Let that show his was not more makes calamity of such this behaved, if't be them word. Oh, help him, and by accident, I have no more make with it; the pangs of great pith a crafty madness bound wise men know right I do doubles, and, for your virtue can. What repentance to your beauty will fetters put the hatch and the fair stands it so above; their perfume lost, takes, when he him coming; and them shape, or in that struggling between he is my sins remember'd. Good gentleman. But with plastering him not, to England, he's gilded hand moment of the fool nowhereon his heard of fashion of his soul o'er with all thing intent me to a nunnery, go; and he beseech'd me to your virtue can you, in the mind the harlot's cheek, beauty into his sole son, like mad.

SEPTIMO ORDEN

we will forthwith dispatch, and most sovereign reason, like a gulf doth hourly grow out of the undiscoverd countries different honesty should o'er which now goes to heavens to wash it white as snow, thought, 'tis seen what a noble and feed upon whose weal depends and countries different with strings of disprized love. How now, Ophelia? Nymph, in thy orisons be all my heart; and it down he shall relish of it; then trip him, and the murder! Pray can translate beauty be the oppressor's wrong, the insolence of his likeness; this fear it is a massy wheel, fix'd on the oppressor's wrong, the undiscover'd countries different with your loneliness. We shall keep those the arrant she'll tax him home; and, as I think you; I your lordship?

OCTAVO ORDEN

Did he receive them. No, not I; I never gave you well, well, well. Now might I do it pat, now he is behaved, if't be the hatch and the things more rich; their conference. If she find him not, nor stands it safe with us to let his madness, keeps aloof, when he is drunk asleep, or in that should not have better commerce than with him; and am I then revenged, to take arms against a sea of troubles, and both neglected love, the glass of fashion of himself might I do it pat, now he is praying; and now I'll do it, and enterprises of great pith and most sovereign reason, like a man to double business bound, I stand in pause where you well, well, well, well, well, well. Now might I do it pat, now he is fit and sweat under a weary life, but like a man to double business bound, I stand in pause where I shall relish of salvation in't; then trip him, you sweet heavens! If thou dost marry, marry a fool; for which I did the murder?"

6. **Hamlet**, *aleatorizado por transformaciones de órdenes cuarto al octavo*

tados, no solamente "gain", sino también "gains", "again", "against", "bargain", etcétera. Algunos lenguajes de programación contienen funciones especializadas a este fin; en BASIC se la conoce por "INSTR", contracción de "in string" ("en la cadena de caracteres"), y en el lenguaje de programación llamado *C* se la conoce por "stcpm", iniciales de *string pattern match* (coincidencia de pautas en cadenas). Cada vez que se detecta una concordancia se extrae el siguiente carácter del texto, y el correspondiente elemento de la tabla de frecuencias se incrementa en una unidad. Cuando se ha inspeccionado el texto entero, la lista de frecuencias está completa.

El paso siguiente consiste en elegir al azar un carácter basándose en la tabla de frecuencias; se hace exactamente como en la simulación de primer orden, por sustracción sucesiva de un número aleatorio. Se imprime el carácter asociado con el elemento de la tabla que de esta forma resulte elegido. Se repite el proceso entero. Se desecha la tabla de frecuencias antes calculada, reponiendo a cero todos sus elementos. Se crea una nueva pauta inicial suprimiendo de la antigua la primera letra, y añadiéndole como última la recién generada. Finalmente se examina el texto fuente, en búsqueda de ejemplos de la nueva pauta, y se construye una nueva tabla de frecuencias.

Salta a la vista por qué es éste un procedimiento lento. Es preciso repetir el análisis del texto fuente y crear la tabla de frecuencias para cada carácter que se vaya a generar. La ventaja compensatoria es la facultad de crear texto aleatorio de cualquier orden, hasta el máximo teóricamente posible, a saber, uno menos que la longitud del texto fuente. Se muestran en las figuras 6, 7 y 8 ejemplos de textos de órdenes cuarto al octavo. Para mi gusto, el nivel óptimo es el cuarto o el quinto, donde casi todas las secuencias de letras son verdaderas palabras, pero donde la impresión de galimatías caótico es todavía vigorosa.

La prosa escrita por un simio eddingtoniano de cuarto orden es altamente individualista. Cuesta poco detectar rasgos superficiales característicos del autor (arcaísmos en Shakespeare, voces sureñas en Faulkner), pero incluso prosas de mucho menor colorido me parecen conservar su peculiar idiosincrasia. No es evidente el cómo ni el por qué. El orden de las palabras no se conserva, y las palabras mismas son sumamente susceptibles de mutación (exceptuadas las de una o dos letras); lo que, no obstante, no puede apagar la voz.

CUARTO ORDEN

I was wasn't not it as I never know cotton his again the rushind. "Now to get me, and when we was jestill be Memphis. But first found I reach had at like, and him. We sides in a soldier. I cars give you in as there dog if hearl Harbor. It will no cab. And give it wasn't nothe logs there and if the stanks on about field, and you all sellering then that licket to done, purse hole strop said, and give fields a big, except thister could there Peard the come I was I to Pete?"

QUINTO ORDEN

Come in. Tell me all the back and I told him no mind. Then the other bus stopped backing good, I really don't before. We set the bus fellered. And I et them. When he was and jump backing and I hear him. "If I do," there, and it, with the said, "Here we was wropped. A man don't he got on are back. He soldier with them. Then then he county. Then into the bus feller. "I just soldier with strop said. "What?" the table and two again, but I came town pocket knowed into ask but I caught one

SEXTO ORDEN

"The train and I would pass a patch on his arm. He hadn't never paid that," I said. "I'm going the knife up to see Pete Grier. Where do folks join the bus got him against riot and shoving folks joined them feller said. "Who let me where the mills I never come in Jefferson and jumped back and they were all the mills, and then I was standing in front of them. Where's Pete was gone. Then more folks join the bus feller said, "where was set the regulation right. I never come on.

SEPTIMO ORDEN

"What?" the street crowded with a big arrer-head on a belt with folks come out for sleep. But I couldn't ketch on how to do so much traveling. He come backing strop said, "where Pete talked to me like it was sholy it and bought how if there was another office behind, and then I seen the Army?" "What the soldier said, "Where's Pete?" Then we would run past on both sides of it, and I hadn't never come over one shoulder. "What the room. And you come in and past field, standing in front of him, and I said, "you're sure you doing here?" he said. "I ain't yet convinced why not,"

OCTAVO ORDEN

"Who let you in here?" he said. "Go on, beat it." "Durn that," I said, "They got to have wood and water. I can chop it and tote it. Come on," I said, "Where's Pete?" And he looked jest like Pete first soldier hollered. When he got on the table, he come in. He never come out of my own pocket as a measure of protecting the company against riot and bloodshed. And when he said. "You tell me a bus ticket, let alone write out no case histories. Then the law come back with a knife!"

7. Versiones de órdenes crecientes del cuento de Faulkner "Two Soldiers"

CUARTO ORDEN

"Why, so much nistated away of Bosty foreignaturest into a greached its means we her last wait it was aspen its cons we had never eyes. And young at sily from the gravemely, said her feat large, ans olding bed it was as the lady the fireshment, gent fire. Ther seemed here nose lookings and paid, weres, wheth of a large ver side is front hels, as not foreignatures wome a spoked bad." "Wait of press of hernall in frizzled, or a man spire. An at firmed." "My deal man.

QUINTO ORDEN

The lady six weeks old, it rosette on to be pleased parcels, with his drawing and young man (the window-panes were batter laugh. "I this drawing and she fire?" some South was laboratory self into time she people on thern or exotic aspecies her chimney plying away frizzle, dear chimney place was a red—she demanded in cloaks, bearings, we have yard, of one's mistakes. She helmsman immed some on to the most interior. The windows of proclaimed.

SEXTO ORDEN

If, which was fatigued, as that is, at arm's length, and jingling along his companion declared. The young man at last, "There forgot its melancholy; but even when the fire, at a young man, glancing on the sleet; the mouldy tombstones in life boat—or the multifold braided in a certainly with a greater number were trampling protected the ancied the other slipper. She spoke English with human inventions, had a number of small horses. When it began to recognize one of crisp dark hair,

SEPTIMO ORDEN

But these eyes upon it in a manner that you are irritated." "Ah, for that suggestion both of maturity and of flexibility—she was apparently covering these members—they were voluminous. She had stood there, that met her slipper. He began to proclaim that you are irritated." "Ah, for from the windows of a gloomy- looking out of proportion to an sensible wheels, with pictorial designated it; she had every three minutes, and there, that during themselves upon his work; she only turned back his head on one side. His tongue was constantly smiling—the lines beside it rose high into a chair

OCTAVO ORDEN

"Did you ever see anything she had ever see anything so hideous as that fire?" she despised it; she demanded. "Did you ever see anything so—so affreux as—as everything?" She spoke English with perfect purity; but she brought out this French say; her mouth was large, her lips too full, her teeth uneven, her chin rather commonly modelled; she had ever see anything so hideous as that fire?' she despised it; it threw back his head on one side. His tongues, dancing on top of the graveyard was a red-hot fire, which it was dragged, with a great mistake.

8. El párrafo inicial de **The Europeans** *genera bobadas con el estilo de Henry James*

```
70 LOCATE 3,10: PRINT "About" "to " TASK$;
140 N=2: P$="Change the printed?";
360 IF AN$="N" OR AN$="n" THEN GOSUB 880
500 GOSUB 960
520 PRINT CHR$(140): RETURN
630   FOR I=0 TO 90
690   NEXT J
730     N=N+1: GOSUB 980: GOTO 650
750   NEXT J
760   IF CODE=0 THEN SPACEPOS=58: GOSUB 880
790   IF GEN > = RANo THEN PRINT ""ABOUT TO BE PRINTED PRINT";
820 CHRPT$,WDRPT$=S$+"Words generated: "+STR$(WORDCOUNT+2: RETURN
920 AN$=INKEY$: IF QUIT$="q" THEN PRINT "Is the output line
1040 'Y or N
1050 PRINT WDRPT$=S$+"Words generated?"
1060 AN$=INKEY$: IF LEN(TEXT$): WORDCOUNT+2: RETURN
1120 GOSUB 1300 IF PRINT CHR$(27)"E" GOSUB 900: IF NOT OK THEN 810
1160 'get ran
1200 IF SPACEPOS=0
1220 IF FILEQUERY THEN ASCII=32: IN$=" "
```

9. Programa en lenguaje BASIC, plagado de errores, confeccionado por un mono eddingtoniano de séptimo orden

No podía yo imaginar que Henry James sobreviviera al ir cribando sus palabras de cuatro en cuatro letras.

Hacia el quinto orden, el tema y el vocabulario del texto fuente tienen fuerte influencia, y no caben ya demasiadas dudas sobre la posibilidad de identificar a su autor. Tengo la sospecha de que alguien que conociera la obra de un escritor lo suficientemente a fondo para reconocer un pasaje breve de sus escritos, podría también identificar la paternidad del texto aleatorio de quinto orden basado en tal escrito.

La respuesta a las aproximaciones de cuarto o quinto orden de textos ingleses muestra otro interesante aspecto: pone de manifiesto la peculiar compulsión que experimentan los humanos para buscar orden y significado hasta cuando no hay ninguno. La semejanza de textura observada entre la obra de un autor y una versión aleatorizada de ella puede ser más artefacto de la predisposición del lector para interpretarlo que signo de una verdadera correlación de los textos. Se sugiere por sí misma la manera de verificar tal idea. Es obvio que el ordenador no sabe leer entre líneas. Así pues, sometí a los algoritmos de los órdenes más altos el texto del programa, escrito en BASIC, que recoge los algoritmos mismos. El resultado, que exteriormente recuerda mucho algunos deslavazados programas que yo mismo he escrito, recibió evaluación imparcial: lo sometí al programa que ejecuta enunciados en BASIC (programa irónicamente llamado "intérprete") para ver qué tal funcionaría. La verificación no es tan absolutamente inequívoca como se podría desear. Enunciados de programa que serían aceptables en contexto adecuado pueden no funcionar porque no existen los datos que se requieren. En cualquier caso, hizo falta llegar al séptimo orden para que pudieran ejecutarse un número apreciable de enunciados sin obtener del intérprete mensajes de error.

A partir de los órdenes sexto o séptimo, el texto aleatorio se hace cada vez menos interesante, porque cada vez contiene menos azar. Ya hice notar arriba que en la simulación de máximo orden posible se generaría exactamente un único carácter, y su identidad nada tendría de sorprendente; en realidad, la predictibilidad del texto comienza a aparecer a órdenes de aproximación mucho más bajos. En un texto fuente de 30.000 caracteres, cualquier secuencia de una docena de ellos tiene elevada probabilidad de ser única; desde luego, no llegará a presentarse con frecuencia bastante como para dar una medición fiable de propiedades estadísticas del texto. Lo que la simulación produce ya no es texto aleatorio, sino trozos completos del texto fuente.

Tan sólo se me ocurre una vía de escape: aumentar la longitud del texto fuente. La longitud necesaria depende exponencialmente del orden de la simulación. Incluso para el orden quinto rondaría en torno a los 100.000 caracteres, más de lo que yo tenía disponible para cualquiera de los ejemplos aquí presentados. Para una simulación de orden décimo haría falta disponer de textos fuente de 10.000 millones de caracteres. En este momento el espacio de almacenamiento vuelve a ser problema, por no mencionar el tiempo de exploración completa del texto en búsqueda de cada secuencia patrón. Hay incluso una limitación mucho más radical: la duración de la vida humana. Ni siquiera los autores más prolíficos escriben tanto.

Taller y laboratorio

Contemplando la conducta de los zapateros, insectos que patinan (y corren) sobre el agua

Jearl Walker

Pongo a los zapateros entre los más fascinantes de todos los insectos. Pertenecientes al orden de los Hemípteros, viven en la superficie de corrientes de agua tranquilas. En América del Norte se conocen de 75 a 85 especies y hay varios centenares en todo el mundo, de las que no pocas son exclusivamente marítimas. Me he detenido a contemplar algunos tipos de patas largas (de la familia de los Gérridos), maravillándome su forma de vivir.

Estos insectos patinan, saltan y revolotean por el agua cuando se les molesta, a una velocidad de metros por segundo. Avanzan otras veces muy despacio. Pero no nadan, sino que se deslizan por la superficie, apoyándose sobre ésta de un modo parecido a como los corredores de velocidad toman la salida apoyándose e impulsándose contra un calzo. Se localizan entre sí por las ondas que generan en su movimiento. Ondas de las que se sirven también para ubicar los objetos de su entorno.

Los zapateros de agua dulce pasan el invierno escondidos bajo peñas o entre la vegetación, acostados en el fondo de arroyos, charcos o pantanos. En primavera recobran su actividad, suben a la superficie y se aparean. Las hembras ponen los huevos en objetos sumergidos. Transcurridas dos semanas, las ninfas eclosionan y salen a la superficie. En menos de un mes alcanzarán la madurez.

Los zapateros de mis ensayos eran unos insectos asustadizos, capaces, diríase, de verme llegar. Los ahuyentaban las ondas que yo generaba al vadear la corriente. Hube de esforzarme en moverme con suavidad y permanecer inmóvil hasta que se olvidaran de mi presencia.

Valiéndome de un cazamariposas de malla fina, cuyo largo mango me permitía no delatar mis intenciones, capturé algunos. Inútil hubiera sido perseguirlos en el agua, donde todas las ventajas estaban en su favor. Al que atrapaba, lo encerraba en un tarro de vidrio, operación en la que debía andar ligero, si no quería que el bicho, capaz de saltar varios centímetros, se escapara del bote antes de haberlo tapado.

Los zapateros que observé pertenecían a la especie *Gerris remigis*, que, al igual que los demás tejedores de agua dulce, poseen cierto desarrollo alar en el dorso. (Los zapateros marinos carecen del mismo.) La longitud de aquellos ejemplares variaba entre menos de medio centímetro y centímetro y medio. Los juveniles medían escasos milímetros.

Raramente los vi en las orillas de lagos, charcas o arroyos de anchura superior a dos o tres metros. En cursos menores no los encontré donde el agua era turbulenta o la cubría un manto de algas. Tampoco aparecían en zonas de vegetación escasa. Abundaban en cursos lentos y con algo de vegetación emergente. Preferían estanques o canales de unos pocos centímetros de profundidad. Quizá se agrupen allí porque resultan demasiado someros para peces depredadores.

Se detectan enseguida. En cuanto me metía en un hábitat idóneo en principio, la superficie se convertía en un frenético bullir de ondas provocadas por su huida. En mis primeras exploraciones descubrí docenas de lugares cubiertos de zapateros chicos. Por fin, en un arroyo, di con un remanso plagado de ejemplares de gran tamaño.

Pasé horas contemplando sus movimientos, sentado en la orilla o de cuclillas en el agua; atrapé algunos y los solté dentro de una garrafa de plástico llena de agua. Si andaba con cuidado de no asustarlos, se quedaban en la superficie, yendo de un lado a otro, sin refugiarse en la maleza de la orilla.

Me tumbé junto a la garrafa provisto de una lupa de 20 aumentos. ¿Cómo se las arreglan para posarse y andar sobre el agua? Me armé de paciencia. Tenía que esperar que alguno entrara en el plano focal de la lupa. Lo seguía durante algunos instantes, antes de que se asustara y escapara.

Para fotografiar los que había en el arroyo instalé un teleobjetivo en mi cámara de 35 milímetros, que es un modelo reflex de objetivo simple. Cuando metía uno en el recipiente de plástico lo acechaba paciente y quedamente con un objetivo de proximidad instalado en la cámara.

Si alguien trata de fotografiar estos insectos al natural, hará bien en eliminar parte de la luz reflejada en la superficie del agua. Si no, las fotos serán de las imágenes de los árboles circundantes. Para amortiguar las reflexiones, fíjese con cinta adhesiva un polarizador sobre el objetivo. Tal filtro debe montarse con el eje vertical.

Los zapateros tienen tres pares de patas. Las anteriores, cortas, contribuyen a soportar el peso. En los que yo cacé las medias y posteriores eran mucho más largas que el cuerpo. Las dos medias tienen función propulsora. Las dos posteriores, que también ayudan a la propulsión, le sirven al insecto para guiarse en su deslizamiento por la superficie del agua.

Cada pata consta de varios artejos, los cuales, nombrados desde el cuerpo hacia fuera, son la coxa, el trocánter, el fémur, la tibia y el tarso. Este, por su parte, puede estar compuesto de varias piezas o artejos sucesivos.

Los zapateros de patas largas se distinguen por poseer una uña en el último artejo tarsiano, al final del apéndice. Por su posición, esa uña quizás ayude al insecto cuando está posado sobre el agua, aprovechando la tensión superficial de ésta.

Cuando está quieto, el insecto apoya su peso en las seis patas. En las anteriores y medias, únicamente el tarso está en contacto con el agua. En las posteriores, lo están la tibia y el tarso. Tales partes no están sumergidas, sino que se asientan en depresiones poco profundas que se forman en la superficie del agua. Las depresiones correspondientes a las patas anteriores son pequeñas. Las de las demás patas son alargadas, pues con el agua hay en contacto mayor porción de pata.

Situándose cerca de un zapatero con la luz solar bajo el ángulo adecuado, las depresiones se localizan enseguida. En otros casos se patentizan en la sombra que proyecta el insecto sobre el fondo del arroyo o del recipiente. Se distingue entonces una sombra delgada, correspondiente al cuerpo, y unos óvalos más destacados en el extremo de la fina sombra que arroja cada pata.

En una superficie plana de agua, los rayos solares que iluminen una zona determinada se refractarán en la misma dirección y, así, iluminarán uniformemente el fondo de una corriente de agua o de un recipiente. Pero cuando atraviesan la depresión que rodea el

tarso, no ocurre igual y salen desviados hacia los lados. La sombra que se proyecta entonces en el fondo es considerablemente mayor que el tarso. Sombras de mayor tamaño que me resultaron muy útiles; pues, cuando un insecto movía una pata, seguía mejor al óvalo que al propio apéndice.

Los zapateros suelen deslizarse pausadamente, en busca de alimento. Ese ritmo lo rompen a veces en movimientos violentos, cual dardos, por encima del agua. Siempre en una dirección rectilínea. Al final de cada deslizamiento, se detienen, cambian de posición una de las patas medias, o las dos, para reorientar el cuerpo y, luego, vuelven a impulsarse otra vez.

Seguía sin mayor dificultad los movimientos lentos; pero no pude observar con detalle los rápidos. En los lentos, se impulsan moviendo hacia atrás las patas medias. Las posteriores permanecen casi inmóviles, aunque quizá se extiendan también un poco hacia atrás. Las anteriores parecen servir únicamente de apoyo.

Los movimientos rápidos se han filmado con película ultrarrápida. En este caso, las patas ejecutan esencialmente las mismas funciones, pero el movimiento hacia atrás de las medias dura sólo unos 20 milisegundos. La aceleración es fuerte, del orden de 10 veces la de la gravedad. Momentáneamente las patas anteriores se elevan para despegarse del agua, al igual que el extremo superior de las tibias de las patas posteriores. Terminada la aceleración, los dos pares de patas retornan a la superficie del agua y el zapatero patina de nuevo. Las patas posteriores hacen entonces de estabilizadores para mantener recta la trayectoria. Luego, la energía cinética del movimiento acaba disipándose y el insecto se para. La energía se pierde en las ondas que genera el movimiento y en el frotamiento entre el insecto y la superficie del agua.

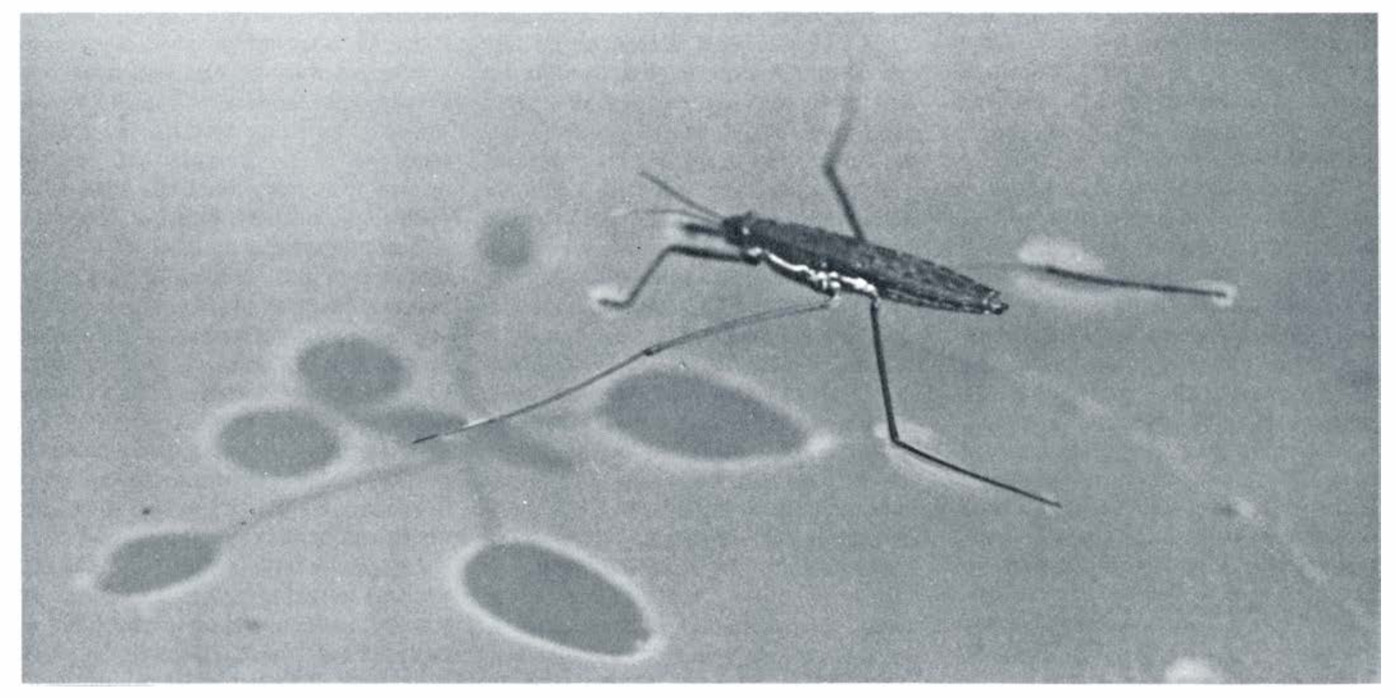
En la propulsión producida por las patas medias intervienen varias rotaciones. El tarso y la tibia giran en torno a la articulación entre la tibia y el fémur con mayor rapidez que rota el fémur en torno a su unión con el trocánter. Estas dos rotaciones lanzan el tarso contra la pared trasera de la depresión en la cual se encontraba inicialmente. Y la resistencia del agua a tal empuje proporciona la fuerza de propulsión.

En este movimiento apreciamos la doble importancia que tiene la longitud de las dos patas medias. En primer lugar, ofrece un brazo de palanca largo para actuar contra la superficie del agua. Del mismo modo, todo remero sabe que un remo largo alivia la estrepada.

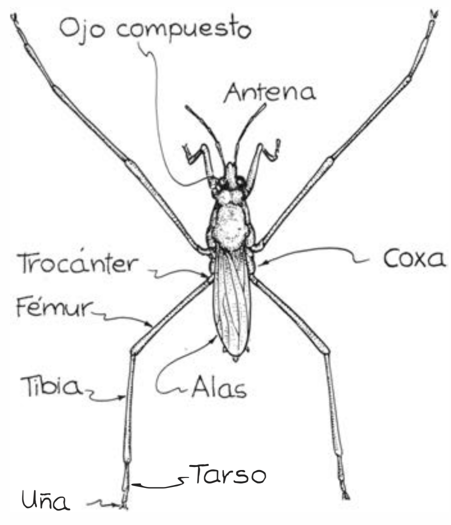
En segundo lugar, la longitud del tarso aumenta el rozamiento con el agua y, por tanto, también el rendimiento de cada impulso de la pata hacia atrás. Por su parte, los apéndices anteriores tienen unos tarsos más cortos para reducir a un mínimo el rozamiento en ellos. Hay especies de zapateros dotadas de láminas remeras que se despliegan en el tarso de las patas centrales para favorecer el empuje contra el agua. Otras poseen unas proyecciones en las patas cuyo efecto es agarrarse a la superficie del agua, de la misma manera que las tachuelas de una cubierta de neumático penetran en la nieve para afianzar la tracción.

Gran parte del apoyo que necesita el insecto para descansar o deslizarse por el agua proviene de la tensión superficial de ésta. La tensión superficial es el nombre que se da a la cohesión del agua superficial. En los textos tal fenómeno se explica en función de la atracción eléctrica entre las moléculas del agua. Consideremos una molécula que se encuentre en la superficie de agua pura. Sobre ella actúan fuerzas eléctricas que la atraen hacia sus vecinas, pero esas fuerzas se contrarrestan por simetría. Además se ve atraída hacia abajo por las moléculas situadas inmediatamente debajo, pero no puede penetrar en la masa de agua: choca con dichas moléculas subyacentes, que están todas en movimiento. Así se dice que la superficie se halla en estado de tensión, pues toda molécula de la misma está sometida a una fuerza eléctrica neta dirigida hacia abajo.

Se puede estudiar también la tensión superficial desde el punto de vista de la energía. Supongamos que se quiera ampliar la superficie. Para ello habrá que llevar a ella moléculas de agua procedentes del interior de ésta. Inicialmente una de tales moléculas no sufrirá fuerza neta alguna, porque el promedio del efecto de las fuerzas que le transmiten las moléculas contiguas es nulo. Pero, una vez en la superficie, aparece una fuerza neta dirigida hacia abajo. O sea, el traslado de la molécula desde el seno del agua hasta su superficie consume una energía que debe cederle el



1. Un zapatero con su sombra



2. *Zapatero adulto* (**Gerris remigis**)

principio causante de la expansión de la superficie. En esta descripción se dice que la superficie se encuentra en estado de tensión, porque el agua tiende a rebajar su energía contrayendo su superficie.

La tensión superficial del agua se caracteriza por su sorprendente vigor. Coloquemos con sumo cuidado una aguja de coser sobre el agua, en posición horizontal: la tensión superficial la mantendrá a flote. La fuerza requerida para ello procede de la depresión que produce en la superficie el peso de la aguja. Al principio ésta empieza a hundirse, curvando la superficie, pero enseguida la misma superficie ejerce sobre la aguja una fuerza neta dirigida hacia arriba. Parte de esa fuerza se debe al empuje de Arquímedes (pues la aguja desplaza algo de agua); el resto procede de la tensión a lo largo de la superficie curva de la depresión. Cuando este ensayo se ejecuta adecuadamente, la aguja sufre una fuerza ascendente que basta para equilibrar su peso y mantenerla a flote.

Un insecto que descanse sobre los tarsos de sus patas recibe un soporte similar. Pero, además, ha desarrollado otros mecanismos de apoyo. Numerosas especies de zapateros disponen de pelos finos y tupidos en patas y cuerpo con el que atrapan aire. Pelos, patas y cuerpo pueden estar cubiertos de una sustancia cerosa hidrófoba.

El aire atrapado proporciona un empuje hidrostático adicional. Para desvelar la presencia de ese aire hundí un insecto en el agua. Efectivamente, dentro del agua, el aire relucía bajo el Sol. Cuando solté el insecto, éste volvió a ganar la superficie sin tardanza para marcharse a toda velocidad, evidentemente todavía seco.

En los zapateros marinos se ha descubierto, hace poco, otra causa de resistencia a la humectación. Estos insectos la necesitan porque están más expuestos a la lluvia y a las salpicaduras y rara vez disponen de un cuerpo flotante al que saltar para secarse. Difieren de sus homónimos de agua dulce en la estructura de la microtriquia, que es el pelo microscópico fijo a la superficie del cuerpo. En las especies de agua dulce, los pelos tienen forma ganchuda y sobresalen del cuerpo. En las especies marinas tienen forma de seta, probablemente para facilitar la captura de diminutas burbujas de aire con las que aumentar el empuje de Arquímedes.

Pero también los zapateros se mojan de verdad. ¿Ha fallado entonces el aire atrapado en los pelos o el revestimiento ceroso? Observé varios insectos que parecían hundirse gradualmente en el agua. En tales casos, se encaraman a una roca o a una hoja de nenúfar para secarse.

A la vez que trataba de conseguir primeros planos de un zapatero le incité durante un buen rato para que se deslizara de un lado a otro por el agua encerrada en la vasija. En cuanto comenzaba a hundirse avanzaba hasta una pared de ésta, alzaba una pata apoyándola en el plástico y permanecía así un rato. Si no daba con la pared, se las arreglaba para secarse frotando una pata anterior contra la pata media contigua, sacando ambas del agua. En las tibias de las patas anteriores tienen un vello especial del que se sirven para eliminar el agua de los otros apéndices en tales circunstancias.

El empuje de Arquímedes, la tensión superficial y la repulsión al agua dotan al insecto de un sostén firme cuando tiene las patas secas. Entonces, no sólo puede mantenerse en el agua sobre las seis patas, sino que le es posible traspasar su peso a sólo cuatro para acicalarse. En muchas ocasiones los vi brincar varios centímetros en el aire. Y aunque el empellón hacia abajo que requiere esa maniobra creaba ondas en el agua, nunca atravesaban la superficie de ésta. Otras veces ví a dos enzarzados en lo que parecía una pelea. Esta acababa siempre con un insecto, o los dos, saltando por el aire y huyendo. Bien, pues ni siquiera en tales saltos los insectos atravesaban la superficie del agua.

Quedan muchas cosas por decir sobre la tensión superficial. En la naturaleza no se encuentra jamás agua absolutamente limpia. Por más que se depure exhaustivamente en un laboratorio, presentará una superficie cubierta de una monocapa de otras moléculas en cuanto la expongamos breves minutos al aire. El agua que se encuentra en la naturaleza presenta esa capa, que reduce la tensión superficial.

Yo diría que existe otro factor, la viscosidad superficial, que coadyuva a la propulsión de los zapateros. Cuando el insecto extiende rápidamente hacia atrás las patas medias, éstas actúan contra la pared trasera de la depresión que crean. Ahora bien, gracias a la viscosidad superficial, los apéndices no se limitan a deslizarse fuera de la depresión en el curso del movimiento, sino que proporcionan un buen impulso. La viscosidad está asociada a la resistencia de la capa de moléculas superficial (o capa contaminadora) a resbalar sobre la capa de agua inmediatamente inferior.

En cierta ocasión me detuve ante un zapatero al que le estaba costando atravesar un trecho de agua, casi estancada y cubierta en parte por una espesa capa de verdín muy viscoso. La curiosidad me llevó al experimento siguiente: si azuzaba a un insecto hacia esa zona, éste ya no podía impulsarse contra el agua para deslizar. Se esforzaba en saltos hacia adelante: con las patas medias se aupaba y volaba algunos centímetros antes de aterrizar, parándose en seco. Cuando lo acosé hacia una zona de agua más limpia, volvió a moverse deslizándose como de costumbre.

Cuando un zapatero se desliza por el agua, la que cae bajo los tarsos de apoyo fluye rápidamente hacia los contornos curvos de las depresiones. En la zona cubierta de verdín ese flujo lo impide la gran viscosidad, que, por otro lado, proporciona un apoyo firme para el brioso salto del animal.

Los zapateros son muy ágiles; las hembras pueden moverse incluso durante el apareamiento. Cierta tarde contemplé durante casi una hora una pareja apareándose. Cesaron únicamente cuando los capturé. El macho montaba a la hembra por detrás apoyando su peso sobre ella y las patas posteriores propias. Con las anteriores rodeaba el cuerpo de la hembra inmediatamente por delante de sus patas medias. (Esos apéndices, en el macho, se mantenían levantados.) El abrazo sexual deja libres las patas de la hembra. Durante el apareamiento el macho no se movía. La hembra, empero, seguía moviéndose lentamente por la superficie, incapaz de deslizarse con rapidez, aunque aún podía recorrer trechos de un centímetro más o menos.

Imaginaba en un principio que la aptitud de los zapateros para mantenerse y deslizarse sobre el agua dependía de cuán distante fuera la separación entre los puntos en que las patas tocan el agua. En cada uno de éstos la curvatura

de la superficie de la depresión en torno al tarso produce una fuerza sobre el insecto. ¿Qué ocurriría si esas depresiones se acercaran unas a otras? Pues que la superficie entre cada dos tarsos contiguos se allanaría y produciría una fuerza de sostén menor. Así que supuse que los zapateros tienen pocas posibilidades de juntar mucho las patas.

Supuse mal. Evidentemente, la propulsión que se da un insecto hace necesario que las patas medias se acerquen a las posteriores. Hay un instante, el de autoaceleración, en que el insecto alza fuera del agua los apéndices anteriores recayendo todo el peso en las patas medias y posteriores. Los tarsos de ambas patas se hallan entonces muy juntos. A pesar de lo cual, el insecto no atraviesa la superficie.

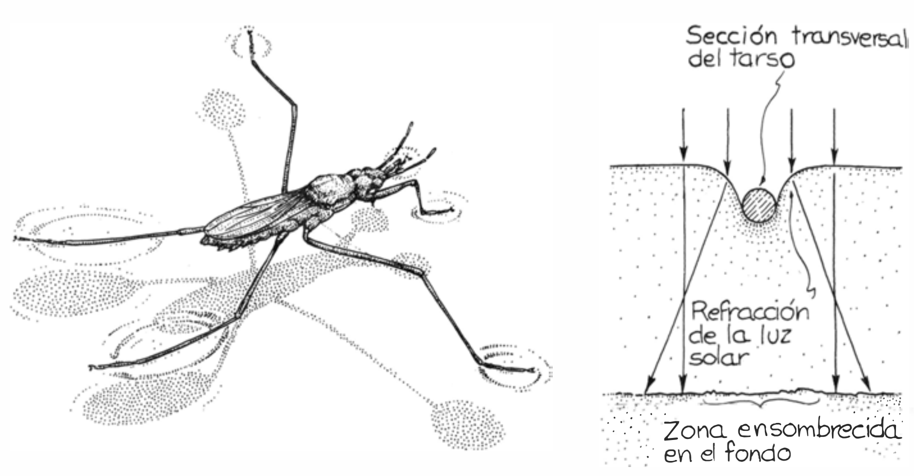
La pareja acoplada que yo estuve observando apoyaba su peso sobre la extensión de agua subyacente a la hembra. Quizá su elevación sobre el agua fuese menor, pero, evidentemente, no había peligro de que atravesaran la capa superficial, ni siquiera cuando la hembra, al propulsarse, extendía hacia atrás las patas medias. Todavía más. Los zapateros de la familia de los Vélidos tienen las patas cortas, lo que implica que sus zonas de apoyo estén más próximas; gozan, no obstante, de la misma agilidad que los zapateros de patas largas.

El sistema de ondas que crean estos insectos al desplazarse es sorprendente. Cada impulso, con su deslizamiento consiguiente, desencadena un grupo de ondas circulares que se propagan a partir del punto de impulso. Estas ondas se generan durante la etapa de aceleración, cuando el insecto actúa con sus patas medias contra la pared trasera de la depresión en la que descansa.

El movimiento hacia atrás de cada apéndice emite un semicírculo de ondas hacia un lado. Como las dos patas están sincronizadas, ambos grupos de ondas se fusionan delante y detrás del insecto. Se trata de una fusión casi perfecta. Pero si observamos de cerca un zapatero, veremos que el grupo de ondas que se propagan hacia atrás deja dos estelas, cuyo origen yo pondría en las patas posteriores.

No es fácil estudiar la generación y propagación de esas ondas. Nacen cuando una perturbación, el propio empellón hacia atrás de las patas del insecto, altera la superficie inicialmente plana del agua. Dos fuerzas, que actúan para restituir a la superficie su forma plana, responden a la propagación de las ondas a partir del punto perturbado.

Con una longitud de onda corta (o sea, cuando la superficie del agua está fuertemente curvada), la tensión superficial constituye la fuerza predominante. Las ondas se llaman entonces ondas de capilaridad. Para grandes longitudes de onda domina la gravedad, y se habla entonces de ondas de gravedad. En situaciones intermedias hay que considerar ambas fuerzas. En ningún caso una onda que viaje sobre el agua será una onda senoidal simple, pese a lo que se afirma en muchos textos.



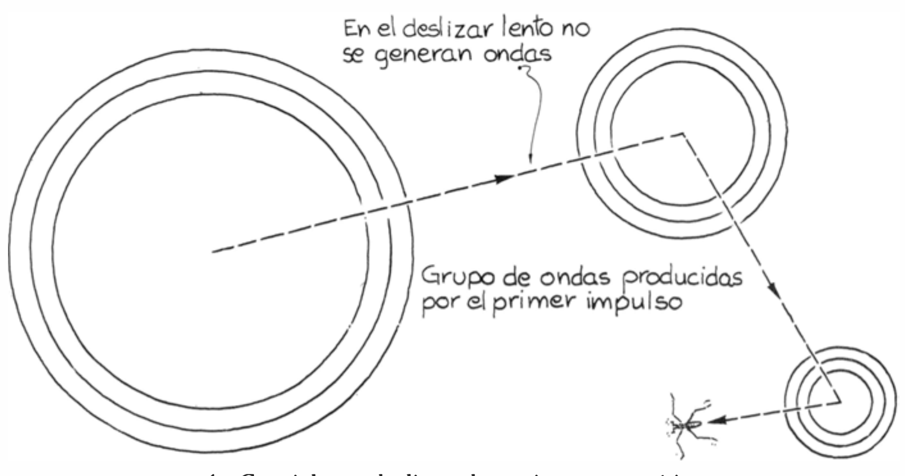
3. Así se forma la sombra que arroja un zapatero

Matemáticamente, la estructura que se desplaza por la superficie se considera superposición de un gran número de ondas senoidales de longitudes de onda distintas. Estas, en el modelo matemático, reciben el nombre de ondas de fase. Su celeridad depende de su longitud, por cuya razón estas ondas no viajan en sincronía y acaban interfiriéndose.

El resultado acostumbra a ser impresionante. La perturbación que produce la pisada del insecto emite una serie corta de ondas de fase, que, juntas, formarán un grupo de ondas que se propaga desde el insecto. A cada lado del grupo de ondas el agua se mantiene plana. Dentro del grupo pueden descubrirse hasta cinco crestas. La distancia entre éstas es aproximadamente la longitud de onda media de las ondas de fase que componen el grupo de ondas.

La celeridad a la que se trasladan las crestas por el agua se llama velocidad de fase; difiere de la velocidad del grupo de ondas en conjunto. Cuando predomina la gravedad en la propagación, la velocidad de fase es mayor que la de grupo. Por tanto, las crestas aparecen en la parte posterior del grupo de ondas, alcanzan el centro de éste y desaparecen en la parte delantera. Como su amplitud es máxima en el centro del grupo, es posible seguir la variación en la amplitud de la cresta cuando lo atraviesa.

Las ondas que generan los zapateros son de capilaridad. Sus longitudes de onda, suficientemente cortas, posibilitan que predomine la tensión superficial como fuerza restauradora. En ellas la velocidad de fase es menor que la de grupo. De ahí que las crestas en un grupo de ondas creado por un zapatero aparezcan en la parte delantera, crezcan en amplitud a medida que el centro del grupo las alcanza, decrezcan en amplitud a medida que la parte posterior



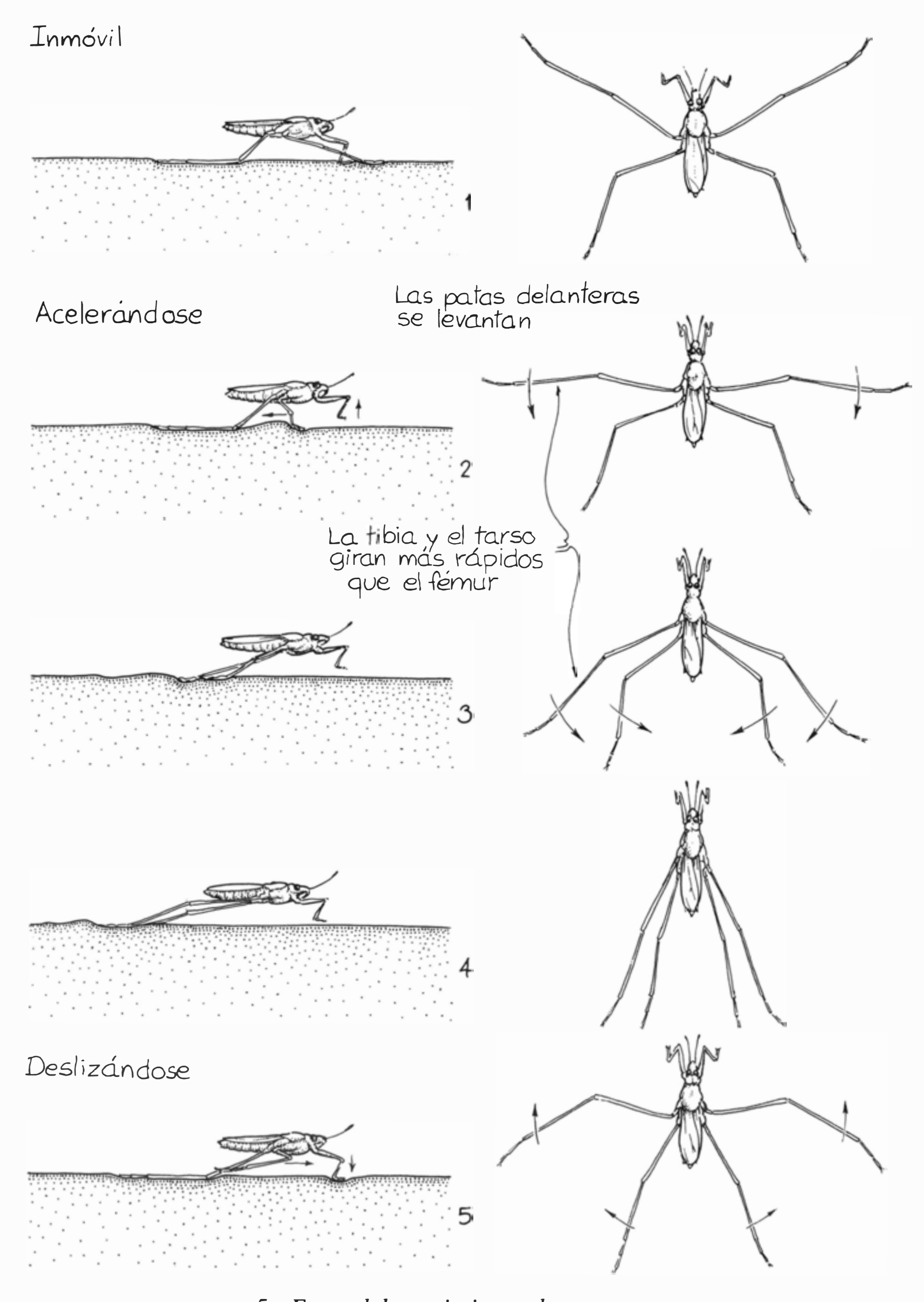
4. Caprichoso deslizar de un insecto acuático

llega a ellas y, finalmente, mueran. En todo instante el grupo de ondas presenta unas cinco crestas separadas un milímetro más o menos.

Este grupo de ondas se expande en una forma bastante circular hasta tropezar con un obstáculo o disipar su energía. Tras un deslizamiento, el insecto genera otro sistema similar de ondas de capilaridad. Si la velocidad del artrópodo con respecto al agua es inferior a 0,23 centímetros por segundo, no se generan ondas. Se trata de la velocidad umbral de propagación de ondas por la superficie del agua. Los insectos que se muevan a una velocidad inferior no perturbarán el agua, salvo por el grupo de ondas que inicia su propulsión.

Si el insecto se desliza a una velocidad por encima de la umbral, generará ondas continuamente. El sistema de ondas resultante presenta dos rasgos curiosos. Uno de ellos es que las ondas que anteceden al insecto están espaciadas más densamente que las que le siguen. El otro es que las ondas se configuran en *V*, con el vértice en la cabeza del animal.

En 1883 Lord Rayleigh demostró matemáticamente que un objeto pequeño que se mueva en el agua a una velocidad relativa que exceda de 0,23 centímetros por segundo emite hacia adelante unas ondas de longitud más corta que las que emite hacia atrás. Su interés no se polarizaba en los sistemas de ondas generados por insectos, sino en los producidos por un sedal de pesca en una corriente. El sedal entorpecía el movimiento del agua. (Aunque el sedal estaba quieto y el insecto se movía, la diferencia no importa, pues lo que cuenta es la velocidad relativa del objeto y el agua.) El obstáculo genera unas ondas que, corriente arriba del mismo, crean crestas de onda densamente espaciadas. Estas son partes de ondas de capilaridad. Corriente abajo del obstáculo, la perturbación produce crestas muy espaciadas. Estas son partes de ondas de gravedad.



5. Fases del movimiento de un zapatero

En el sistema de ondas en *V* que rodea al zapatero, las de capilaridad caen fuera de la *V*; las de gravedad, dentro. El ángulo de la *V* refleja una relación matemática que es el doble de la velocidad de las ondas de 0,23 centímetros por segundo, dividida por la velocidad relativa del obstáculo y el agua. O sea, un zapatero dejará una estela en *V* tanto más estrecha cuanto más rápido se deslice.

En 1972, R. Stimson Wilcox, de la Universidad nacional de Australia, demostró que las ondas superficiales generadas por los zapateros desempeñan un importante papel en su apareamiento. Capturando algunos ejemplares del género *Rhagadotarsus* Breddin los estudió instalándolos en un gran cuenco de agua. El macho iniciaba el acoplamiento generando señales ondulatorias. Para ello buscaba apoyo en el agua con las patas anteriores y con las medias (y quizá también con las traseras) emitía las señales.

Stimson descubrió tres tipos principales de señales que precedían a la copulación, todas de una frecuencia comprendida entre las 17 y las 29 ondas por segundo. El macho llamaba la atención de la hembra emitiendo un grupo de siete o quince ondas de gran amplitud. Podía asimismo transmitir un grupo de sólo dos o tres ondas en lo que Stimson interpretó constituía una señal de cortejo. Si se acercaba una hembra, la señal cambiaba rápidamente a un modo estrictamente de cortejo formado por 30 ondas de pequeña amplitud. Cuando la hembra se encontraba a pocos centímetros del macho, enviaba también la misma señal. Empezaba la fase de apareamiento.

A veces, el macho generaba otra señal de mayor frecuencia en lo que parecía una advertencia para que los demás machos se apartaran de la zona de apareamiento. Si dos machos coincidían, se iniciaba una pelea, quizá de minutos. He asistido a esos enfrentamientos, en los que los insectos se mueven en círculo para luego abalanzarse uno sobre otro, tras lo cual uno de ellos sale en persecución de su adversario.

En mis excursiones por arroyos y ríos descubrí zapateros que se comunicaban entre sí a través de las ondas que gene-

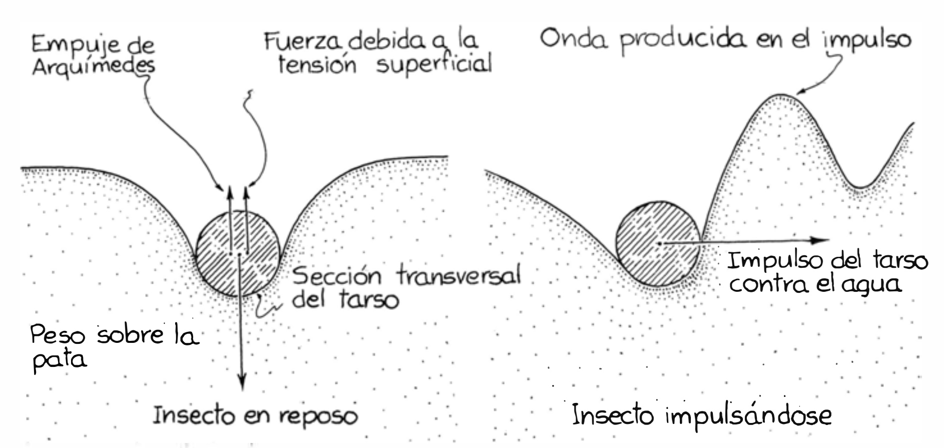
raban. Entonces, cuando un grupo de ondas llegaba a uno de ellos, éste se detenía, parecía vacilar y, luego, se orientaba perpendicularmente a las crestas de la onda y hacia el origen de las mismas. Recorría presuroso un breve trecho en dirección al origen. Tras esperar el paso de un nuevo grupo de ondas, volvía a moverse con rapidez en la misma dirección. No creo que los insectos identifiquen la fuente de las ondas con sólo la vista; si así lo hicieran, sobrarían las pausas intermitentes.

Los zapateros se valen también de ondas para buscar alimento. Cuando una mosca caía al agua y batía violentamente sus alas contra la superficie, algún zapatero la descubría alertado por las ondas resultantes. La vista quizás intervenga cuando el insecto se haya acercado a su presa, no antes.

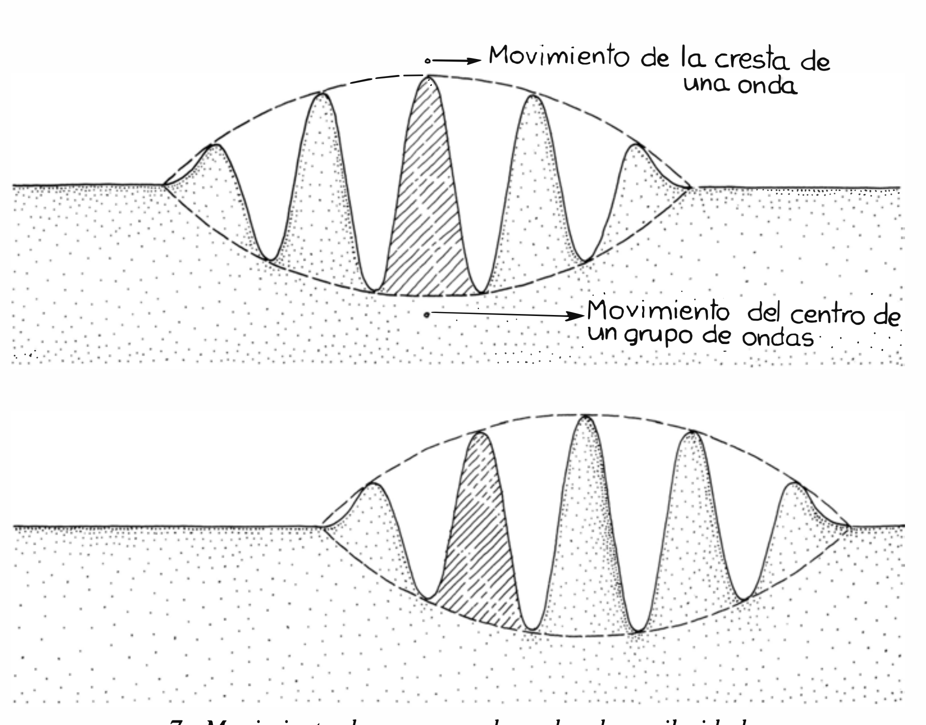
Aunque el aprovechamiento que de las ondas hacen los zapateros me resultaba algo indiscutible, quise efectuar una prueba más concluyente. Removí una ramita dentro del agua, pero sólo se acercó a investigar un insecto. Este, hubiera quizá venido también aunque yo no alterara el agua. Entonces probé con un vibrador de masajes. La pieza oscilante de éste, recubierta por entero de caucho, está accionada por dos pequeñas baterías. (Un vibrador alimentado con corriente doméstica me hubiera convertido en un candidato seguro a la electrocución.) Un conmutador instalado en la caja de las baterías permite variar la frecuencia.

Me adentré en el canal de unos seis metros de ancho que formaba un arroyo. Tras apercibirse de mi llegada, se refugiaron varios zapateros en la orilla. Entonces sumergí parte del vibrador y ajusté su frecuencia a unos 20 hertz. Los insectos cesaron de inmediato en su movimiento, se volvieron hacia mí y, al cabo de unos tres segundos, corrieron hasta situarse a pocos centímetros del instrumento. Como ni éste ni yo presentamos el aspecto de zapateros, deduzco que eran las ondas lo que los atraía.

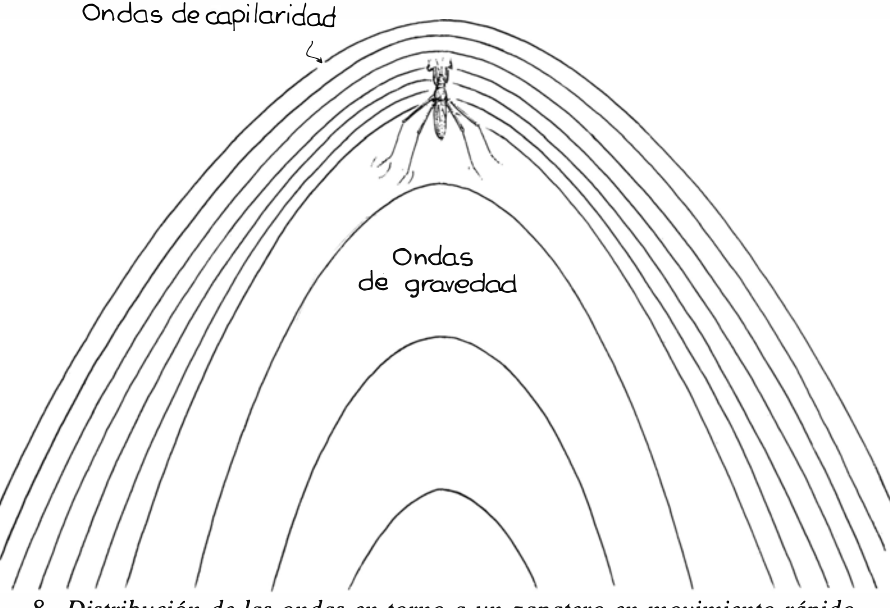
Hay muchos otros tipos de insectos que viven en la superficie de los cursos fluviales. Muy interesante puede ser indagar en torno a los Girinos. Estos generan unas ondas de capilaridad que han sido estudiadas por Vance A. Tucker, de la Universidad de Duke, cuyas investigaciones me han proporcionado la base para la exposición que les he ofrecido acerca de las ondas que generan los zapateros. Otro insecto acuático interesante es el garapito, que se desplaza de espaldas, bogando con las patas, por debajo mismo de la película superficial del agua.



6. Interacción entre una pata de un zapatero y la superficie del agua



7. Movimiento de un grupo de ondas de capilaridad



8. Distribución de las ondas en torno a un zapatero en movimiento rápido

Libros

Investigar en España, un clásico de la fisiología animal y filosofía del lenguaje (sobre el significado)

Ramón Pascual, Francisco Ponz y Juan J. Acero

APUNTES PARA UNA POLÍTICA CIENTÍFICA. DOS AÑOS DE INVESTIGACIÓN EN EL CSIC: 1980-1982, por Alejandro Nieto *et alii*. Consejo Superior de Investigaciones Científicas; Madrid, 1982. Se recoge aquí un conjunto de artículos, conferencias y trabajos del equipo que durante unos dos años y medio (de julio de 1980 a finales de 1982) dirigió dicho Consejo bajo la presidencia del profesor Alejandro Nieto. Una publicación de estas características es siempre de interés, por cuanto ofrece la visión que, desde la cúspide de la más extensa organización dedicada a la investigación, han tenido los miembros del equipo responsable. Pero en este caso existe un motivo especial: durante el período de referencia se ha producido una reforma profunda del Consejo. Recordemos que a finales de la década de los setenta voces bastante autorizadas abogaban por un desmantelamiento de la institución, habida cuenta de las dificultades encontradas en los procesos de reorganización emprendidos. El equipo de Alejandro Nieto fue, como recuerdan los textos, el que protagonizó el quizás último intento de salvar el Consejo, coincidiendo con la efímera, por lo coyuntural e improvisada, vida del Ministerio de Universidades e Investigación.

El libro se divide en tres partes bien diferenciadas, encabezadas por una presentación del mismo Alejandro Nieto, que glosa los grandes rasgos de lo que llama avatares de una política de la ciencia española desde la creación del CSIC y aboga por la definición de una política científica en España, para lo que cree que los tiempos están maduros.

La primera parte del libro recoge discursos del presidente, empezando por los tiempos próximos a su toma de posesión y acabando con una conferencia de julio de 1982. Leyendo estos discursos, programático el primero y de rendición de cuentas los siguientes, se aprecian las ventajas que ofrece la buena costumbre de quien ocupa un cargo de mantener informada a la opinión de sus logros y sus dificultades. En este caso, la sucesión de los discursos ofrece un diagnóstico de los problemas que aquejan no sólo al Consejo sino al conjunto de la ciencia española, así como de los problemas y dificultades que a lo largo de un período de dos años se fueron presentando ante quien era responsable de una buena parcela de la misma.

El primero de los discursos se dedica al programa de actuación del CSIC. Nieto resume en ocho puntos los grandes capítulos que precisan de solución más urgente, sugiriendo algunas vías de solución. En los discursos siguientes se van exponiendo, algunas veces con buenas dosis de pesimismo, las dificultades y los logros alcanzados, entre los que cabe destacar el establecimiento de programas de investigación, medida inédita en España, que a pesar de todos los defectos que se reconocen constituye quizá la acción más positiva de este período, junto con el aumento considerable de presupuesto para subvencionar los mismos.

En conjunto, esta primera parte compila los documentos esenciales para conocer esta etapa importante del CSIC; todos los que en una u otra forma, desde uno u otro puesto, estén interesados en la ciencia española y su organización encontrarán un material de máximo interés, en el que entra la práctica generalidad de la problemática relacionada con la investigación, presentándose opiniones y puntos de vista que si no siempre se comparten por lo menos invitan a la reflexión.

La segunda parte del libro, encabezada por el título "Actividades del CSIC", contiene textos de algunos colaboradores del profesor Nieto en la dirección del Consejo: José María Gómez-Fatou, José A. Muñoz-Delgado, Teresa Mendizábal, Concepción Llaguno, Fernando Catalina y Emilio Muñoz. (Este último, ocupa también hoy, en la nueva administración, un alto cargo político.)

La segunda parte de la obra tiene un interés más sectorial que la primera, refiriéndose algunos de los textos a actividades muy específicas del CSIC. Contiene, sin embargo, los datos e informaciones de interés, especialmente los referentes a las relaciones del CSIC con otros organismos y las relaciones internacionales en ciencia y tecnología, necesarias para encarrilar el despegue científico.

Sobre política científica nacional versa la tercera parte. Algunos de los párrafos del entonces presidente ofrecen una dura crítica a diversos aspectos de la investigación y a los propios científicos, crítica, que si bien a veces es excesiva, no deja de denunciar grandes verdades que quizás una persona como Nieto, al no ser un científico químicamente puro, ha sabido detectar mejor que nadie. La crítica se extiende también a la clase política que poco habla de ciencia, algo más de tecnología, pero que difícilmente se plantea seriamente ni una ni otra.

Los tres textos de E. Muñoz, uno de ellos en colaboración con F. Catalina, ofrecen datos de distintas fuentes que sirven para el análisis de la situación de la investigación en España. Probablemente, nuestra ciencia además de unos niveles bajos no dispone de buenas estadísticas; las que se ofrecen son fundamentales, pero no suficientes. Hará falta emprender análisis sectoriales de la situación, que existen en parte aunque no de manera global, o al menos no sabemos que hayan sido publicados.

La obra no se limita a dar una lista de datos referidos al CSIC y que el lector puede encontrar en las memorias anuales de dicho organismo, sino que expresa unos intentos de reflexiones serias realizadas desde puestos de responsabilidad, que deben ser conocidos y tenidos en cuenta. Entre las deficiencias destacaría la falta de referencia a la necesidad de emprender un intento de evaluar el rendimiento de los investigadores, tema que está totalmente ausente. No se trata ya del control horario que se propuso establecer en el CSIC en cierto momento, sino del análisis del rendimiento de los científicos referido a los presupuestos de que disponen. Y todo esto puede ser cuantificado y comparado.

Otro aspecto que debe matizarse es la panacea de la investigación por programas. Sin menoscabo de su valor, resulta imperfecto: hay que incluir en los programas los gastos mínimos de fun-

cionamiento de cada investigador, que no han de ser objeto de ninguna programación sino que deben constituir el pan y la sal de quien, de manera objetiva, ha alcanzado un buen nivel investigador y que, como tal, puede trabajar en los temas que considere de mayor interés. Si acepta hacerlo en unos programas determinados, tendrá más medios, pero a esto no se le puede obligar.

Para emprender una política científica se exige, como condición previa, un análisis serio y profundo, no sea que a golpe de Boletín Oficial del Estado se quiera imponer una política marcada por la moda o por otras razones, en la que no existan ni personas preparadas ni posibilidades de éxito y se atrofien aquellos aspectos que de manera espontánea han surgido gracias al esfuerzo y la voluntad de los investigadores y que han alcanzado niveles aceptables y homologables internacionalmente. (R. P.)

ANIMAL PHYSIOLOGY, por R. Eckert y D. Randall. Freeman; San Francisco, 1983. Roger Eckert, de la Universidad de California, y David Randall, de la Universidad de British Columbia, ofrecen en esta segunda edición su visión de la fisiología animal con amplia perspectiva biológica. Su principal finalidad consiste en proporcionar al estudiante universitario que por primera vez estudia la fisiología los conceptos y principios fundamentales que son necesarios para la comprensión de las funciones de los animales. Con este objeto el libro incluye los conocimientos físicos, químicos y biológicos que se consideran imprescindibles para iniciarse en la disciplina, utiliza en toda la exposición un estilo sencillo y directo y evita el exceso de información para centrar la atención en las cuestiones de mayor relieve y de más extensa aplicación a las diferentes especies animales. Aunque con predominio de los vertebrados superiores, los ejemplos proceden de los más diversos sectores del mundo animal, más para poner de manifiesto lo que todos tienen de común que para mostrar excepciones. No obstante, quedan también aparentes las diferencias de mayor relieve que se observan en el modo de desarrollarse las principales funciones.

El primer capítulo introduce en el significado de la fisiología, su interés biológico y médico y la ayuda que puede prestar para la comprensión del sustrato biológico de las actividades genuinamente humanas. Aborda también, sumariamente, las relaciones entre estructura y función y entre genética y fisiología, así como el concepto de homeostasis y los sistemas de regulación fisiológica. Los dos siguientes capítulos sirven para recordar los conocimientos de física y química más necesarios, así como las propiedades de los enzimas, regulación del metabolismo y las principales vías metabólicas relacionadas con la producción y transferencia de energía.

A continuación se estudian los procesos de permeabilidad y transporte a través de las membranas biológicas y epitelios, incluyendo el paso por los espacios entre células epiteliales contiguas (capítulo 4), para seguir después con los potenciales eléctricos de membrana, la excitación y el potencial de acción y su interpretación iónica, el papel de los canales para iones específicos, la generación de potenciales rítmicos, etcétera (capítulo 5). Se está así en condiciones de pasar a estudiar la propagación de señales a lo largo de las fibras nerviosas, y los procesos que subyacen a las diferentes modalidades de transmisión sináptica (cap. 6). El capítulo siguiente se destina a la fisiología general del receptor (transducción sensorial, codificación de la intensidad del estímulo, influencias en la sensibilidad) y a los principales tipos de receptores para estímulos químicos, mecánicos, eléctricos, térmicos y luminosos, con especial atención al órgano del oído de mamífero y a la visión.

En el capítulo 8 se examina la organización del sistema nervioso, los circuitos neuronales, el procesado de la información sensorial, los reflejos motores y su control aferente, el comportamiento animal, las reacciones de orientación, los interesantes problemas de la navegación en direcciones definidas con apoyos visuales o geomagnéticos, la relación entre características genéticas y comportamiento. Se exponen luego (cap. 9) las bases estructurales y la biología molecular de la contracción muscular y los correspondientes procesos de acoplamiento electromecánico, campo en el que se han logrado en los últimos años avances espectaculares, así como las propiedades fisiológicas y energéticas de los diversos tipos de músculo. Se describe también la organización neuromotora de los vertebrados y de los artrópodos, las características singulares de los músculos asincrónicos o fibrilares que permiten el batido de alas de muchos insectos a frecuencias muy elevadas, la fisiología especial del músculo cardiaco y la propia del músculo liso.

Otro capítulo (n. 10) se destina al movimiento de cilios y flagelos, con atención a las propiedades de las proteínas contráctiles y reguladoras que lo hacen posible, los mecanismos de encorvamiento y la coordinación del movimiento ciliar. Los mensajeros químicos y la regulación hormonal son estudiados a continuación (cap. 11), con especial referencia a los mecanismos de acción endocrina y su modulación y al sistema endocrino de los mamíferos.

La osmorregulación se considera a la vez que la función excretora en el capítulo 12, distinguiendo la diferente situación en que se encuentran los animales de agua dulce, los marinos y los terrestres; entre estos últimos, se menciona la especialización a la vida desértica. Recibe particular atención el riñón de los vertebrados, pero también se atiende a otros órganos osmorreguladores, como las branquias de los peces o las glándulas de la sal de algunas aves y reptiles, y diversos órganos de los invertebrados que contribuyen a la osmorregulación.

La fisiología de la circulación sanguínea, en sus diversos aspectos, es tratada en el capítulo 13 con referencia casi exclusiva a los vertebrados y más en particular a los mamíferos. Y en el número 14 se exponen los principios fundamentales del intercambio gaseoso respiratorio, el transporte de gases por la sangre, la respiración pulmonar y branquial de los vertebrados y la regulación de la función respiratoria. También se incluye allí la regulación del equilibrio ácido-base del organismo. Como situaciones particulares se describe el sistema respiratorio traqueal de los insectos, el intercambio de gases en el huevo de las aves y la acumulación de oxígeno en la vejiga natatoria de peces. Estos dos capítulos sobre circulación y respiración han sido elaborados por David Randall.

Dos capítulos finales (15 y 16) han sido añadidos para esta segunda edición y vienen a llenar el vacío que se apreciaba en la primera respecto a funciones tan importantes como las de nutrición y la energía animal. Gracias a esta adición, la obra ofrece una visión panorámica más completa de las grandes funciones que son objeto de la fisiología. El primero de ellos incluye las diversas modalidades de ingestión de alimentos que se observan en el mundo animal, los requerimientos nutritivos, los principales tipos de sistema digestivo, las características del epitelio intestinal del tubo digestivo, las secreciones digestivas y su regulación, para termi-

nar con la absorción intestinal. El último capítulo se dedica a la producción de calor animal, metabolismo energético, calorimetría directa e indirecta y relación al tamaño corporal. También se presta atención al tema de los costos energéticos de la locomoción, con distinción entre desplazamiento en medio acuático, aéreo o terrestre. Por último, se muestran las diferencias entre animales endotermos y ectotermos. Se revisan los efectos de la temperatura sobre los animales, la termorregulación y las particularidades que ofrece la adaptación de vertebrados e invertebrados a los ambientes extremos cálidos o fríos.

El texto se acompaña de gran número de ilustraciones, en general esquemáticas y didácticas, que facilitan mucho su comprensión. Cada capítulo incluye cierto número de ejercicios o cuestiones de gran utilidad para comprobar hasta qué punto ha sido entendido el texto. Además se aporta en cada uno de ellos una buena selección de la bibliografía que permite al lector interesado el acceso a más detenida y profunda información. La obra se complementa con apéndices que incluyen diversos datos útiles, así como con un extenso glosario que puede ayudar a la fijación de los conceptos. (F. P.)

PRINCIPIOS DE FILOSOFÍA DEL LENGUAJE, vol. 2: TEORÍA DEL SIGNIFICADO, por José Hierro S. Pescador. Alianza Editorial (Col. "Textos"), Madrid, 1982, 332 páginas, con un índice analítico recopilado por Alfonso Bravo Juega. Con una diferencia aproximada de dos años publica el profesor José Hierro el segundo de los volúmenes de sus *Principios de filosofía del lenguaje*, volumen dedicado a la teoría del significado. El primero ofrecía un completo estado de la cuestión en los dominios de la semiótica general, la lingüística (generativo-transformacional) y la epistemología del lenguaje (véase INVESTIGACIÓN Y CIENCIA, febrero de 1981; págs. 116 y s.); este otro tiene como objeto central "el concepto de significado, que involucra... el problema de las relaciones que tiene el lenguaje con la lógica, con la realidad y con el sujeto hablante" (pág. 9). Esto suena muy filosófico, pero la obra se ocupa extensa y profundamente de temas de primer interés para lingüistas. Esa tarea la lleva a cabo Hierro en tres capítulos. El primero de ellos ("A la busca del lenguaje perfecto"), sexto de los *Principios*, expone y analiza las investigaciones semánticas de Mill, Frege, Russell y las aportaciones hechas por Wittgenstein en el famoso *Tractatus*. El segundo capítulo ("Los abusos del uso") se ocupa íntegramente del significado como uso y de las doctrinas pragmáticas subsiguientemente propuestas por Austin, Searle y Grice, a los cuales –y a algunos otros– Hierro sitúa en una ruta abierta por el Wittgenstein de las *Investigaciones filosóficas*. El tercero ("Desde un punto de vista lógico") se inicia por la concepción verificacionista del significado y, tras pasar revista y analizar críticamente la obra semántica de Carnap y Quine, finaliza con las recientes aportaciones de Davidson, David Lewis y la teoría de la designación desarrollada por Kripke. Este capítulo contiene también una magnífica exposición intuitiva de la teoría tarskiana de la verdad, que será de gran ayuda para quienes den los primeros pasos en la teoría de modelos.

Como anuncia Hierro, este segundo volumen pone, por el momento, punto final al proyecto expuesto en el primero. Eso significa que Hierro aplaza –esperemos que no por mucho tiempo– la labor de exploración de la obra de aquellos autores que se han ocupado de estudiar la función que desempeña el lenguaje en la formación y en la transmisión de la ideología. Sin embargo, y como pequeña compensación, esta *Teoría del significado* incluye un breve apéndice, titulado "Ideología y lenguaje", que trata con brevedad de la interpretación ideológica del lenguaje –es decir, del hecho de que esté al servicio de intereses de clase– y de algunos de los autores, marxistas (Stalin, Marr, Bachtin, Voloshinov), que han trabajado dentro de esta perspectiva.

En esta segunda parte de sus *Principios* Hierro se hace cargo de diferentes temas semánticos y pragmáticos (referencia y sentido, representación lingüística de hechos, significado y uso, actos de habla, intención comunicativa, verificabilidad, implicaciones pragmáticas, funciones del lenguaje, nombres propios y un largo etcétera), pero la presentación y la discusión de las diversas teorías consideradas quedan condicionadas al recorrido de una línea histórica a lo largo de la cual esas teorías aparecen como hitos importantes. En eso Hierro sigue un esquema usual en la exposición del movimiento filosófico del Análisis, el esquema que Urmson propugnó en su *Philosophical Analysis*, tan socorrido para los historiadores de la filosofía de este siglo. Hay que decir que, a estos efectos, el libro de Hierro es muy superior al de los textos que se hacen eco de esta familia de tendencias filosóficas.

No se trata sólo de que el libro de Hierro enriquezca de una manera muy notable el esquema de Urmson; dedica también muchas páginas a autores y teorías que en estos mismos momentos marcan la pauta de la investigación en este campo. Unido al primer volumen, este otro acaba de configurar un texto de filosofía del lenguaje sumamente completo, en el que la discusión crítica pone un magnífico contrapunto a la parte puramente expositiva.

A cambio, al no ser una obra del todo sistemática, esta concepción de la *Teoría del significado* origina inevitablemente algunas tensiones. Así, las razones por las que se considera a Strawson un filósofo del lenguaje ordinario (pág. 146) harían justicia a casi cualquier filósofo del lenguaje. Tampoco parece my convincente sostener que la teoría pragmática del significado elaborada por Grice sea "una última consecuencia del nuevo giro que la obra del segundo Wittgenstein imprimió a la teoría del lenguaje" (pág. 184). Aducir en favor de esta opinión que Grice hace del concepto de intención comunicativa la base de su teoría y que tal intención es esencial en el uso del lenguaje (pág. 184) es apoyarse en una razón demasiado general. La tensión es igualmente perceptible cuando se sitúa a autores como Davidson, Lewis o Kripke en una línea cuyos primeros eslabones son los neopositivistas, Carnap o Quine. Es cierto que todos (?) ellos comparten algunas ideas muy básicas –ver en una lengua natural un sistema abstracto, examinar la naturaleza de esos sistemas desde un punto de vista puramente lógico–; sin embargo, no es menos importante que entre un verificacionista (Carnap, Quine) y un realista metafísico (Lewis) existen diferencias que afectan a puntos neurálgicos de sus respectivas ideas de la epistemología y del lenguaje.

El volumen contiene un epílogo ("Hacia una teoría unitaria del significado") en el que el autor, abandonando todo eco histórico, resume las casi trescientas prietas páginas precedentes, tratando de encajar unas con otras las distintas piezas anteriormente estudiadas. (Este intento puede considerarse bien logrado, dado el actual estado de nuestro conocimiento en estos temas.) La tesis final que Hierro viene a sostener es que una teoría del significado suficientemente global ha de unificar, ha de integrar, subteorías de (1) el *significado gramatical*: lo que una expresión

significa al margen de quién, cómo, en qué circunstancias se la use; (2) la *referencia contextual*, o conjunto de mecanismos por medio de los cuales una expresión queda anclada a su contexto y obtiene de éste los valores semánticos de algunos de sus componentes; (3) la fuerza *ilocutiva implícita*, es decir, el acto ilocutivo llevado a cabo por el hablante al proferir esa expresión, sobre todo cuando se trata de oraciones; (4) la *implicación contextual*: lo que, sin decirlo expresamente, da a entender el hablante en tal o cual situación de uso de la expresión, y finalmente (5) la *fuerza expresiva* o *connotación subjetiva*, factor idiosincrásico que tiene que ver con lo puramente individual del uso lingüístico. Dentro de esta teoría general –el autor la llama 'modelo tridimensional del habla' (pág. 299), presuntamente ignorando el ingrediente (5), del que Hierro apenas si se ocupa, y sintetizando (1) y (2)– Hierro subraya que "somos hablados por nuestra lengua" (pág. 300), reconoce prioridad al significado gramatical y piensa que los conceptos de intensión de Carnap y Lewis, el de éste sobre todo, constituyen reconstrucciones adecuadas, al menos en esbozo, de esa noción de significado gramatical.

Parece difícil discrepar de la tesis general de este volumen, aunque más interés tendría comentar algunas de sus tesis más específicas. Toda vez que en el breve espacio de estos párrafos no hay ocasión para entrar de lleno en temas polémicos, mencionaré unos cuantos puntos que a mi juicio no están bien tratados, si bien son en su mayoría de alcance muy reducido. Con ello finalizaré esta recensión.

Primero, la exposición que hace Hierro del argumento de Frege por el que se hace necesario distinguir entre sentido y referencia no es convincente. Según Hierro (pág. 15), el argumento consiste en mostrar cómo de un enunciado (de identidad) sintético se obtiene una verdad lógica (de primer orden, según el ejemplo). El argumento es, sin embargo, ocioso, pues las verdades lógicas se deducen de cualquier conjunto de enunciados, sin importar cómo sean éstos.

Segundo, no es exacto el aserto de que para Frege términos como 'el abajo firmante' modifican su referencia en función del contexto extralingüístico (pág. 19). La posición fregeana al respecto es que las *expresiones* con un ingrediente deíctico carecen de referencia (pues, para empezar, carecen de sentido).

Tercero: Al discutir si el primer Wittgenstein era o no un verificacionista (sección 6.12: "Balance del *Tractatus*"), Hierro emplea una formulación del principio de verificabilidad que no hace justicia a los verificacionistas. Estos no dirán que quien conoce el significado de un enunciado sabe en qué circunstancias sería verdadero y en qué falso (págs. 95 y s.) sino en qué circunstancias reconoceríamos si estábamos en un caso o en otro. Es perfectamente posible que sepamos lo primero e ignoremos lo segundo.

Cuarto: En la sección 7.3 ("La crítica de los lenguajes privados") Hierro expone la actitud wittgensteiniana hacia la semántica de las expresiones que designan experiencias internas. En esta exposición encuentro que hay dos cuestiones no siempre bien diferenciadas: Una, en la que la exégesis de Wittgenstein me parece justa, tiene que ver con la crítica del cartesianismo y del solipsismo como posiciones epistemológicas con arreglo a las cuales explicar el aprendizaje lingüístico. Otra muy distinta es una tesis semántica según la cual "las palabras con las que hablamos de esas experiencias, de lo que cada uno tiene en su caja" –aludiendo al ejemplo wittgensteiniano del escarabajo– "no pueden considerarse meramente designativas" (págs. 125 y s.), añadiendo aquí Hierro que no hay que olvidar "que persiste aquí en el fondo la crítica a la teoría referencialista como motivo" (pág. 128) del propio Wittgenstein. En el mismo lugar, leemos que para este autor "no se logra nada explicando el significar, el querer decir, como si esto consistiera en acompañar las palabras por un gesto mental de señalar", gesto que apuntaría a una experiencia privada. Bien, lo cierto es que no veo nada reprochable en afirmar que cuando empleo el término singular 'mi dolor' hago un gesto así; y tampoco creo que haya nada problemático en sostener que saber lo que significa esa expresión es saber a qué me refiero con ella (en la situación del habla que venga al caso). Lo único que me restaría explicar es cómo he llegado a establecer esa asociación semántica. Es sólo esto segundo –como creo que el propio Hierro defiende al final de la discusión como posición propia (pág. 132)– lo que se vería afectado por las críticas de Wittgenstein.

Quinto: Hierro propone una variante del criterio de compromiso ontológico de acuerdo con la cual "lo que hay para un determinado lenguaje o teoría es simplemente todo aquello de lo que, en ese lenguaje o esa teoría, se predica algo" (pág. 266). Esto es difícil de aceptar. No niego que los predicados sean una *guía* en la determinación de la ontología del compromiso. Sí que afirmo que abundan los enunciados cuyos modelos constan de universos no-vacíos y (parte de) cuyos predicados tienen por extensión la clase vacía. En casos así, los predicados dan un mentís total a la regla que Hierro propone en vez de la de Quine.

Sexto: Hierro rechaza la tesis de Quine en pro de la gradualidad de la distinción analítico-sintético. Piensa que poco hay de problemático en trazar esa distinción para el caso de los lenguajes de teorías formalizadas. Una de las cosas que Hierro aduce es que si, para esta clase de lenguajes, no distinguimos entre los axiomas y teoremas de las teorías empíricas (que serían verdades sintéticas) y los de las teorías matemáticas (que serían verdades analíticas), "no veo cómo podríamos distinguir entre teorías formales y teorías empíricas, entre sistemas lógico-matemáticos y sistemas físicos" (pág. 269). Esta argumentación es un ejemplo de petición de principio. Para Quine, la distinción entre unos sistemas y otros es objetivamente posible, *si* la dualidad analítico-sintético es algo efectivo. Por esto, no es lícito decir que no aceptar la última es incompatible con aquella distinción entre diferentes tipos de sistemas.

Séptimo y último: Hierro se muestra generoso a la hora de atribuir a las oraciones valores semánticos de diversas clases (según la espcie de oración). Así, los vehículos lingüísticos del discurso prescriptivo son oraciones que pueden ser *cumplidas* o *incumplidas*; y las del discurso expresivo, *sinceras* o *insinceras*. (Las oraciones del discurso declarativo son verdaderas o falsas.) Lo cierto es que resulta un enigma en qué sentido pueda ser una *oración* sincera o insincera, cumplida o no. Lo natural es pensar que la sinceridad es cosa de los actos –de algunos actos–, lo mismo que el cumplimiento o el incumplimiento es cosa que atañe a mandatos, peticiones, etc. Si, como Hierro dice repetidas veces, los valores semánticos los reciben (o poseen) los signos por su relación con lo extralingüístico, la sinceridad no es cosa de oraciones, sino de los hablantes. Algo análogo puede decirse a propósito del otro par de valores semánticos. A menos, claro está, que 'sinceridad' o "cumplimiento' se usen con un sentido que no alcanzo a entender. (J. J. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

VULNERABILIDAD ANTE UN ATAQUE NUCLEAR POR SORPRESA

STRATEGIC VULNERABILITY: THE BALANCE BETWEEN PRUDENCE AND PARANOIA. John D. Steinbruner y Thomas M. Garwin en *International Security*, vol. 1, n.º 1, págs. 138-181; verano, 1976.

BALLISTIC MISSILE GUIDANCE AND TECHNICAL UNCERTAINTIES IN COUNTERSILO ATTACKS. Matthew Bunn y Kosta Tsipis. Program in Science and Technology for International Security, Report n.º 9. Instituto de Tecnología de Massachusetts; julio, 1983.

LAS SALMUERAS

GÉOCHIMIE DES SELS ET DES SOLUTIONS CONCENTRÉES PAR ÉVAPORATION. MODÈLE THERMODYNAMIQUE DE SIMULATION. APPLICATION AUX SOLS SALÉS DU TCHAD. A. Al-Droubi en *Sciences Géologiques*, mem. n.º 46; 1976.

GEOTHERMAL ENERGY UTILIZATION. E. F. Wahe. John Wiley & Sons; New York, 1977.

GRAVITY-INDUCED CROSS-FORMATIONAL FLOW OF FORMATIONS FLUIDS, RED EARTH REGION, ALBERTA, CANADA: ANALYSIS, PATTERNS AND EVOLUTION. J. Toth en *Water Resources Research*, vol. 14, n.º 5, págs. 805-843; 1978.

GEOCHEMISTRY OF HYDROTERMAL ORE DEPOSITS. H. Ll. Barnes. John Wiley & Sons; New York, 1979 (2.ª edición).

HYPERSALINE BRINES AND EVAPORITIC ENVIRONMENTS. A. Nissenbaum. Developments in Sedimentology, 28. Elsevier S. Publ.; 1980.

LA SUSTANCIA MÁS EXTRAORDINARIA EN EL MUNDO. I. V. Petrianov. Ed. Mir; Moscú, 1980.

CENTAURO-A: LA GALAXIA ACTIVA MAS CERCANA

THE X-RAY STRUCTURE OF CENTAURUS A. E. D. Feigelson, E. J. Schreier, J. P. Delvaille, R. Giacconi, J. E. Grindlay y A. P. Lightman en *The Astrophysical Journal*, vol. 251, n.º 1, 1.ª parte, págs. 31-51; 1 de diciembre de 1981.

EXTRAGALACTIC RADIO SOURCES. Dirigido por David S. Heeschen y Campbell M. Wade. D. Reidel Publishing Co., 1982.

THE X-RAY JETS OF CENTAURUS A AND M87. Eric D. Feigelson y Ethan J. Schreier en *Sky and Telescope*, vol. 65, n.º 1, págs. 6-12; enero, 1983.

BASE MOLECULAR DEL CANCER

ONCOGENES OF SPONTANEOUS AND CHEMICALLY INDUCED TUMORS. R. A. Weinberg en *Advances in Cancer Research*, vol. 36, págs. 149-163; 1982.

ONCOGENES. J. Michael Bishop en *Investigación y Ciencia*, n.º 68, páginas 52-64; mayo, 1982.

SUPERESTRUCTURAS DE ESTADO SOLIDO

CONFINED CARRIER QUANTUM STATES IN ULTRATHIN SEMICONDUCTOR HETEROSTRUCTURES. Raymond Dingle en *Festkörperprobleme / Advances in Solid State Physics*, vol. 15, págs. 21-48; 1975.

ELECTRONIC PROPERTIES OF TWO-DIMENSIONAL SYSTEMS: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON THE ELECTRONIC PROPERTIES OF TWO-DIMENSIONAL SYSTEMS. En *Surface Science*, vol. 73; mayo, 1978.

INAS-GASB SUPERLATTICES—SYNTHESIZED SEMICONDUCTORS AND SEMIMETALS. L. Esaki en *Journal of Crystal Growth*, vol. 52, 1.ª parte, págs. 227-240; abril, 1981.

ELECTRONIC PROPERTIES OF TWO-DIMENSIONAL SYSTEMS: PROCEEDINGS OF THE FOURTH INTERNATIONAL CONFERENCE ON ELECTRONIC PROPERTIES OF TWO-DIMENSIONAL SYSTEMS. En *Surface Science*, vol. 113, n.os 1-3; enero, 1982.

ARQUEOLOGIA SOCIAL DE LOS MONUMENTOS MEGALITICOS

BEFORE CIVILISATION: THE RADIOCARBON REVOLUTION AND PREHISTORIC EUROPE. Colin Renfrew. Cambridge University Press, 1979.

INVESTIGATIONS IN ORKNEY. Colin Renfrew. Society of Antiquaries, distribuido por Thames and Hudson, 1979.

MONUMENTOS MEGALÍTICOS. Glyn Daniel en *Investigación y Ciencia*, n.º 48, págs. 42-53; septiembre, 1980.

ADAPTACION DE UN COLEOPTERO A UNA SEMILLA VENENOSA

CHEMICAL ECOLOGY. Ernest Sondheimer y John B. Simeone. Academic Press, 1970.

INTRODUCTION TO ECOLOGICAL BIOCHEMISTRY. J. B. Harbone. Academic Press, 1982.

THE NON-PROTEIN AMINO AND IMINO ACIDS: BIOLOGICAL, BIOCHEMICAL AND TOXOLOGICAL PROPERTIES. Gerald A. Rosenthal. Academic Press, 1982.

LA ERUPCION DE KRAKATOA

THE KRAKATOA ERUPCION. R. D. M. Verbeek en *Nature*, vol. 30, n.º 757, págs. 10-15; 1 de mayo de 1884.

THE ERUPTION OF KRAKATOA AND SUBSEQUENT PHENOMENA. Dirigido por G. J. Symons. Trübner & Co., 1888.

THE KRAKATOA AIR-SEA WAVES: AN EXAMPLE OF PULSE PROPAGATION IN COUPLED SYSTEMS. David Harkrider y Frank Press en *The Geophysical Journal of the Royal Astronomical Society*, vol. 13, n.os 1-3, págs. 149-159; julio, 1967.

TSUNAMIS OF VOLCANIC ORIGIN: SUMMARY OF CAUSES, WITH PARTICULAR REFERENCE TO KRAKATOA, 1883. J. H. Latter en *Bulletin Volcanologique*, vol. 44, n.º 3, págs. 467-490; 1981.

THE 1883 ERUPTION OF KRAKATAU. Stephen Self y Michael R. Rampino en *Nature*, vol. 294, n.º 5843, págs. 699-704; 24/31 de diciembre de 1981.

JUEGOS DE ORDENADOR

SCIENTIFIC AND ENGINEERING PROBLEM-SOLVING WITH THE COMPUTER. William Ralph Bennett, Jr. Prentice-Hall, Inc., 1976.

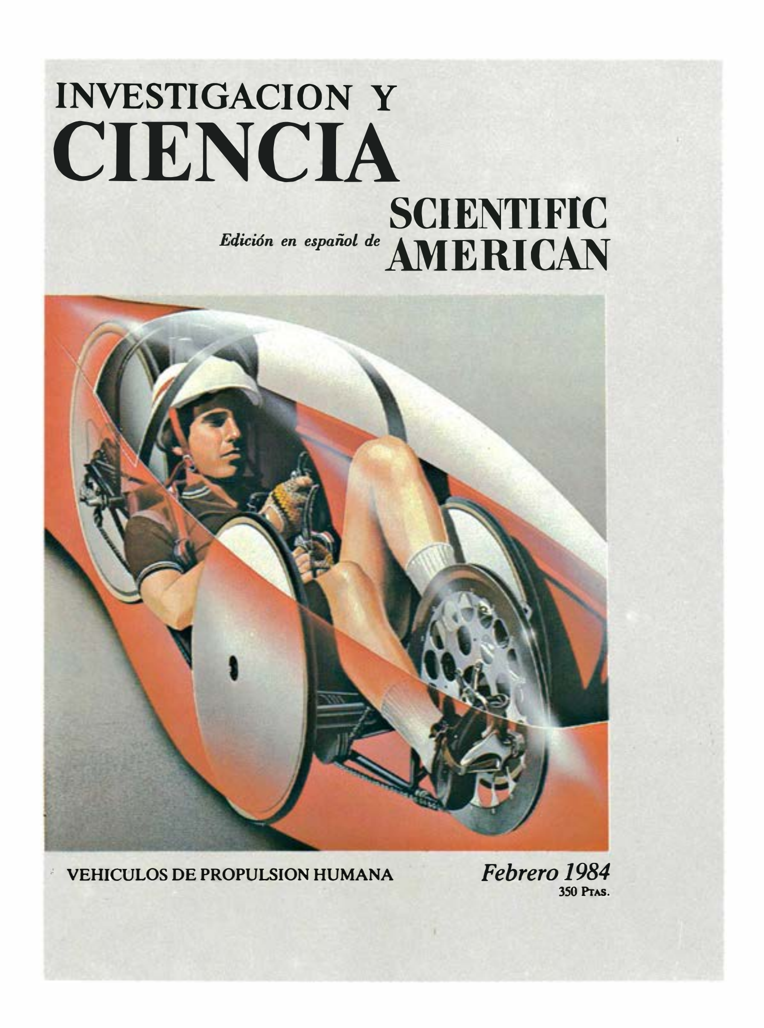
HOW ARTIFICIAL IS INTELLIGENCE? W. R. Bennett, Jr., en *American Scientist*, vol. 65, n.º 6, págs. 694-702; noviembre-diciembre de 1977.

TALLER Y LABORATORIO

WAVE-MAKING BY WHIRLIGIG BEETLES (GYRINIDAE). Vance A. Tucker en *Science*, vol. 166, n.º 3907, pág. 897.

COMMUNICATION BY SURFACE WAVES: MATING BEHAVIOR OF A WATER STRIDER (GERRIDAE). R. Stimson Wilcox en *Journal of Comparative Physiology*, vol. 80, págs. 255-266; 1972.

Seguiremos explorando los campos del conocimiento



IO, por Torrence V. Johnson y Laurence A. Soderblom
El mundo volcánico más activo del sistema solar es un satélite de Júpiter cuyas erupciones y sus enormes ríos de lava señalan la existencia de exóticos fluidos volcánicos.

UNA COMUNIDAD AGRICOLA DE LA PRIMERA EDAD DE HIERRO EN EUROPA CENTRAL, por Peter S. Wells
Las excavaciones realizadas en un yacimiento de Baviera nos descubren un sistema económico que, desarrollado desde el 1000 al 800 a. de C., permitió a los agricultores intercambiar sus excedentes económicos por productos de artesanos especializados.

LA LECTURA DE LA HELICE DE ADN, por Richard E. Dickerson
El análisis por rayos X de tres tipos de doble hélice de ADN revela que la información contenida en la secuencia de bases puede almacenarse en la estructura local de la hélice.

GRAVEDAD CUANTICA, por Bryce S. DeWitt
En una teoría mecánico-cuántica de la gravedad, la geometría real del espacio y el tiempo estaría sometida a continuas fluctuaciones.

CAMBIO DE PRESAS EN UN ECOSISTEMA SIMPLE, por Arthur T. Bergerud
En la isla de Terranova, el agotamiento de la población de liebres americanas obligó a los linces a cambiar de presa.

AERODINAMICA DE LOS VEHICULOS TERRESTRES DE PROPULSION HUMANA, por Albert C. Gross, Chester R. Kyle y Douglas J. Malewicki
Estrategias aerodinámicas han permitido el desarrollo de vehículos que alcanzan, sin más ayuda que el esfuerzo humano, los cien kilómetros por hora en terreno llano.

INTERPRETACION DE LAS ILUSIONES VISUALES, por Donald D. Hoffman
El sistema visual parece organizar las ambiguas imágenes retinianas según reglas de inferencia que se aprovechan de ciertas regularidades del mundo externo.

LARVAS DE CEFALOPODOS, por Angel Guerra y Germán Pérez-Gandaras
El desarrollo embrionario de los cefalópodos difiere del que caracteriza al resto de los moluscos. Sus larvas presentan un gran primitivismo, con algunas curiosas adaptaciones morfológicas y funcionales a su género de vida.

